Concepciones de la Referencia

Eleonora Orlando

eudeba



enciclopedia lógica

Director de colección
Eduardo Alejandro Barrio

Concepciones de la referencia

Eleonora Orlando





Eudeba
Universidad de Buenos Aires

1ª edición: agosto de 1999



Av. Rivadavia 1571/73 (1033)
Tel: 4383-8025 / Fax: 4383-2202
www. eudeba.com.ar

Diseño de tapa: Juan Cruz Gonella
Corrección y composición general: Eudeba

ISBN 950-23-0979-0
Impreso en Argentina.
Hecho el depósito que establece la ley 11.723

Enciclopedia Lógica

En la actualidad, la lógica se ha convertido en una disciplina con un desarrollo complejo. Debajo de la búsqueda de una respuesta a los clásicos interrogantes sobre lo que caracteriza el buen razonar o aquello que distingue a las verdades de la lógica del resto, se han ido abriendo paso nuevos y profundos desafíos. De todos ellos, gran parte contiene un notable contenido filosófico. ¿Es posible que exista un lenguaje capaz de contener todas las verdades? ¿Existe el conjunto de todos los conjuntos? ¿Cómo somos capaces de hablar y de sacar consecuencias lógicas acerca de objetos que no existen? Esta colección intenta explorar en éstas y muchas otras preguntas con el propósito de despertar en el lector la perplejidad que caracteriza a los filósofos. La idea es brindar un análisis claro, preciso, detallado y actualizado de los principales conceptos de la lógica contemporánea. La tarea emprendida no es del todo fácil de realizar, dada la complejidad y a la vez el estado embrionario de muchos de esos conceptos. Cabe destacar que el carácter intrincado del área estudiada ha determinado la ausencia casi total de libros que cubran la totalidad del espectro de sus problemas, por lo que la colección propuesta resulta novedosa.

La serie está dirigida a un universo heterogéneo de lectores. Lo está para todos aquellos a quienes les interese la lógica y en especial, los problemas filosóficos que a ella subyacen (sean ellos iniciados o

no en la vasta problemática que en ella se trata). Las obras están escritas por especialistas, pero intentan no presuponer en los lectores conocimientos sofisticados sobre los temas tratados. Por ello, los textos pueden ser utilizados como apoyo para cursos introductorios. La colección también está dirigida a los filósofos, lingüistas, psicólogos o especialistas en ciencias cognitivas preocupados por los problemas que plantea la lógica. En especial, a aquellos interesados en los problemas involucrados en la argumentación. En suma, está dirigida a todos aquellos lectores que tengan interés en conocer la variada problemática que en la actualidad subyace a la lógica formal.

En este volumen, Eleonora Orlando analiza la problemática relacionada con el valor semántico de las expresiones del lenguaje de la lógica que supuestamente representan individuos y conjuntos de individuos. En su concepción de la lógica, el tema está estrechamente relacionado con la explicación del valor semántico de los nombres y términos generales en nuestro lenguaje natural. La autora defiende implícitamente una continuidad conceptual entre los lenguajes artificiales y naturales. Situado en este marco, el libro contiene una interesante y muy inteligente respuesta al principal desafío planteado por quienes descreen que los lenguajes tengan una función representativa fundamental. ¿Cómo se conectan nuestras palabras con el mundo? Un entramado de relaciones causales y aspectos descriptivos constituye la clave de la respuesta que Eleonora Orlando defiende con originalidad, claridad y profundidad.

Quiero agradecer a las autoridades de EUDEBA, quienes confiaron en mí para la realización de este proyecto, y a todos aquellos que me apoyaron con su interés, su trabajo, sus críticas y especialmente su afecto.

Eduardo Alejandro Barrio
Director de la colección

ÍNDICE

AGRADECIMIENTOS 9

INTRODUCCIÓN 11

CAPÍTULO 1 :LA CONCEPCIÓN DESCRIPCIONAL
Y EL MODELO HISTÓRICO-CAUSAL 29
1. La concepción descripcional 30
1.1. Los nombres propios: Frege vs. Mill 30
1.2. La teoría de las descripciones de Russell 42
1.3. Los problemas de la teoría clásica 49
1.4. Las teorías "cúmulo" 50
2. Kripke o el origen del modelo causal 56
2.1. Designadores rígidos y causalidad 56
2.1.1. La propuesta negativa: nombres y términos generales vs. descripciones 57
2.1.1.1. La tesis de los designadores rígidos y el argumento modal-epistemológico 58
2.1.1.2. El argumento semántico 67
2.1.2. La propuesta positiva: un esbozo del modelo causal 71
2.2. Los problemas de una teoría causal de la referencia directa 75
2.2.1. El problema del respecto 75
2.2.2. El problema del sentido 82

CAPÍTULO 2: LAS TEORÍAS MIXTAS 91
1. Putnam y las teorías del doble factor 92
1.1. El aporte de Putnam 92
1.1.1. En torno a la Tierra Gemela 92
1.1.2. Una nueva dimensión epistémica: el estereotipo 97
1.2. Dos tradiciones para las teorías del doble factor 102
1.2.1. Sentidos, roles conceptuales y referentes 103
1.2.2. Internalismo y externalismo 108
1.3. Un análisis crítico 113
1.3.1. ¿Teoría descriptiva o teoría normativa? 113
1.3.1.1. El problema de la evidencia 114
1.3.1.2. La cuestión normativa 116
1.3.2. El holismo 120
2. La teoría descriptivo-causal de Devitt y Sterelny 132
2.1. Las tesis principales 129
2.1.1. Los sentidos como tipos de cadenas designativas 129
2.1.2. El rol de las descripciones 134
2.2. El análisis crítico 136
2.2.1. La respuesta al problema del sentido 137
2.2.2. La respuesta al problema del respecto 143

CAPÍTULO 3: LAS TEORÍAS CAUSALES NO HISTÓRICAS 151
1. La teoría covariacional de Fodor 153
1.1. ¿Es fatal la ambigüedad? 153
1.2. La respuesta de Fodor 158
1.3. La crítica de la teoría covariacional 160
1.3.1. La cláusula ceteris paribus 160
1.3.2. El problema de Quine 164
1.3.3. Discriminación, idealización y verificacionismo 167
1.3.4. La estrategia de exportación 174
2. El enfoque teleológico-causal: la teoría de Millikan 180
2.1. Una explicación puramente teleológica 180
2.2. La solución del problema de los positivos falsos 185
2.3. En defensa del concepto de función biológica:
una respuesta a Fodor 190
2.4. La objeción panglossianista y la restricción
del alcance de la teoría 195

CONCLUSIÓN 203

BIBLIOGRAFÍA 209

Agradecimientos

Este libro es el fruto de varios años dedicados a la investigación. Agradezco por tanto a la Universidad de Buenos Aires por las becas que me fueron otorgadas entre los años 1988 y 1993. Asimismo, fue muy importante mi participación en diversos proyectos de investigación UBACyT sobre temas de semántica y filosofía de la mente, todos ellos radicados en el Instituto de Filosofía de la Facultad de Filosofía y Letras. Quiero mencionar, además, que este trabajo fue escrito mediante un subsidio de la Fundación Antorchas.

Distintas personas han influido en el proceso de elaboración. Ante todo, quiero agradecer especialmente a Alberto Moretti, mi director de becas y de tesis doctoral. Él fue quien me inició en el estudio de la filosofía del lenguaje y la semántica filosófica, quien inspiró muchas de las ideas aquí presentes. Pero, fundamentalmente, fue quien me dio los mejores ejemplos del ejercicio de la capacidad argumentativa y del espíritu crítico al servicio de una idea, que es lo que caracteriza a la vida profesional del filósofo. Ha sido un privilegio para mí contar con él como guía y maestro.

Mis estudios en el Departamento de Filosofía de la Universidad de Maryland me han permitido estar en contacto con algunos profesores de esa institución. Michael Devitt ha leído y comentado distintos trabajos en los que este libro se ha basado. Por lo demás, sus

seminarios sobre semántica, sus artículos y libros, así como las diversas discusiones mantenidas con él sobre estos temas, han influido enormemente sobre mi manera de encararlos. Quiero expresarle, por ello, un profundo reconocimiento. Georges Rey ha leído y realizado agudas observaciones a mi crítica a la teoría del significado propuesta por Fodor, incluida en el tercer capítulo del libro. El recuerdo de su auténtica pasión por la discusión racional es para mí un estímulo permanente.

Con algunas personas he mantenido discusiones sobre los temas aquí tratados que han enriquecido notablemente mi acercamiento a ellos. Quiero mencionar a Manuel García-Carpintero, Jarek Gryz, Richard Hanley, Plinio Junqueira-Smith, Eric Lormand y Monique Roelofs. Mis reuniones de discusión periódicas con Juan Comesaña, Federico Penelas y Carolina Sartorio han sido una instancia sumamente valiosa para la elaboración de este trabajo.

Finalmente, quiero destinar a Eduardo Barrio un reconocimiento muy especial. El estímulo de su inteligencia y el apoyo de su amistad han sido, y son, para mí, fundamentales y decisivos.

Introducción

Este trabajo está centrado en el análisis del concepto de referencia, esto es, la relación semántica entre los signos descriptivos del lenguaje y los objetos del mundo.[1] Más específicamente, a lo largo de los tres capítulos que lo integran, me ocupo de examinar críticamente las distintas teorías referencialistas del significado que se han presentado, es decir, aquellas teorías que dan un rol central a la referencia en la explicación del significado de los signos en cuestión.[2] La tesis propuesta como consecuencia de este examen es que

1. Es útil aclarar que utilizo la palabra "objeto" en un sentido general, como término genérico, como sinónimo de "entidad" o "cosa" -de modo tal de abarcar distintas categorías específicas, como las de individuos, clases de individuos, propiedades, relaciones o cualquier otra categoría ontológica con la que uno quiera comprometerse. No presupongo ni quiero defender ninguna ontología en particular. (Téngase en cuenta que éste es, fundamentalmente, un libro de semántica y no de ontología).
2. Vale la pena destacar que en ocasiones el adjetivo "referencialista" o "referencial" aplicado a la teoría semántica tiene otra acepción, más restringida, según la cual una teoría referencialista del significado es aquélla que explica el significado solamente en términos del objeto referido (véase, por ejemplo, Alston 1964: cap. 1, pp. 28-29). En este trabajo, sólo se usa la expresión con el sentido amplio consignado en el texto principal.

tal explicación requiere la combinación de factores provenientes de distintas concepciones acerca de la referencia: más específicamente, sostengo que, a fin de explicar el significado del lenguaje natural y del pensamiento o lenguaje mental, es preciso sostener una teoría mixta o ecléctica de la referencia.

Considero conveniente hacer ciertas aclaraciones iniciales. En primer lugar, los signos descriptivos de los que me ocupo en este trabajo son los nombres propios (tales como "Juan" o "Varsovia"), las descripciones definidas ("El autor de *Crimen y Castigo*" o "La protagonista de *Triple Traición*") y los términos generales referidos a tipos o clases naturales, los cuales se subdividen a su vez en nombres contables ("tigre", "átomo") y términos de masa ("agua", "oro"); no examinaré por tanto las explicaciones ofrecidas para las descripciones indefinidas ("una mesa redonda") ni para los términos generales de tipos o clases artificiales ("martillo", "velero"), así como tampoco las teorías propuestas para expresiones indicadoras, esto es, pronombres personales ("ella"), pronombres y adjetivos demostrativos ("éste", "este"), adverbios de lugar y tiempo ("aquí", "ahora"), etc.

En segundo lugar, quiero destacar que la teoría propuesta comprende tanto al significado lingüístico como al significado mental. Cabe tener en cuenta que, si bien las primeras teorías de la referencia ofrecidas se referían sólo al lenguaje natural, esto es, eran teorías linguosemánticas, en la actualidad, muchas de ellas se refieren tanto al lenguaje natural como al pensamiento (o lenguaje mental o lenguaje del pensamiento) y, en ocasiones, sólo a éste último, es decir, son teorías psicosemánticas. En mi opinión, lenguaje y pensamiento son dos instancias profundamente interrelacionadas; más específicamente, suscribo plenamente la hipótesis de la existencia de un lenguaje del pensamiento, pero, a diferencia de autores como Fodor, considero que se trata del mismo lenguaje natural que hablamos o de algo muy cercano a éste. Desde mi punto de vista, dado que las representaciones mentales o conceptos no son sino palabras internalizadas, no hay razones para creer que la habilidad conceptual no esté profundamente relacionada con la habilidad lingüística y dependa en parte de ella. De este modo, creo que las consideraciones linguosemánticas no pueden estar divorciadas de

consideraciones psicosemánticas y viceversa. De ahí el carácter mixto de la teoría propuesta.[3]

En tercer lugar, me interesa señalar que suscribo plenamente la tesis de que la teoría semántica debe tener un carácter empírico: la razón de ello es que considero que la semántica o teoría del significado debe aspirar a convertirse en una ciencia empírica, tanto como la psicología, la lingüística o la economía, y su metodología debe incluir entonces el recurso a la experiencia como único método de justificación.

En este punto, cabe entonces preguntarse por la pertinencia de la inclusión de este libro en la Enciclopedia Lógica, esto es, una colección de trabajos dedicados fundamentalmente al estudio filosófico de los conceptos centrales de la lógica contemporánea. Más específicamente, cabe preguntarse cómo se relaciona el estudio empírico del significado del lenguaje natural y del pensamiento con la reflexión sobre la lógica, esto es, una de las ciencias formales por excelencia. Es oportuno entonces hacer ciertas aclaraciones ulteriores. Los nombres propios del lenguaje natural (y, de acuerdo con lo dicho más arriba, del pensamiento) se corresponden con los términos singulares de los lenguajes formales de los que se ocupa la lógica, mientras que los términos generales tienen sus análogos formales en los predicados. De este modo, por un lado, los nombres propios funcionan semánticamente de manera análoga a como lo hacen los términos singulares: a cada uno se le asigna un individuo específico de un cierto dominio (el mundo entero y el dominio recorrido por las variables respectivamente); por otro, el funcionamiento semántico de los términos generales es semejante al de los predicados: a cada uno se le asigna un conjunto determinado de individuos de un cierto dominio. Por consiguiente, el estudio del funciona-

3. Cabe aclarar entonces que, fuera del contexto de análisis de teorías específicas acerca del lenguaje natural o del lenguaje mental en particular, utilizaré las expresiones "expresión", "signo", "término" y "representación" de modo intercambiable, para referirme no sólo a las *palabras* del lenguaje natural sino también a los *conceptos* del lenguaje mental. Del mismo modo, "oración" servirá para hacer referencia no sólo a las oraciones del lenguaje natural sino también a las del pensamiento.

miento semántico de los nombres propios y los términos generales resulta claramente pertinente para dirimir cuestiones acerca de la interpretación formal de los términos singulares y los predicados respectivamente. En este punto, cabe citar a Susan Haack:

> *Se han usado opiniones rivales sobre cómo entender los nombres propios en los lenguajes naturales para apoyar propuestas alternativas sobre la interpretación formal de los términos singulares en los cálculos menos sencillos, por ejemplo, el cálculo modal.*[4]

El fundamento de la mencionada analogía semántica reside en la estrecha relación existente entre la argumentación informal desarrollada en el lenguaje natural y los argumentos formales de los que se ocupa fundamentalmente la lógica.[5] En síntesis, dada la analogía semántica en cuestión, el estudio de, por ejemplo, las concepciones alternativas de la relación referencial que vincula a los nombres con los individuos del mundo resulta sumamente pertinente para poder adoptar una posición determinada acerca de la interpretación formal adecuada de los términos singulares, esto es, un problema que compete indudablemente a la filosofía de la lógica. De este modo, es posible afirmar que el propósito de este libro ha sido examinar y evaluar las distintas teorías que en definitiva pretenden dar cuenta de la relación, propia de todo sistema de lógica, de asignación de un individuo (o conjunto de individuos) a un término dado, así como de sugerir una posible teoría al respecto. En términos más simples, he intentado esbozar una respuesta a la siguiente pregunta: ¿qué quiere decir que a un término del lenguaje se le asigne un individuo (o una clase de individuos) del universo del discurso? ¿En qué consiste exactamente la mencionada relación de asignación entre términos, por un lado, e individuos, por otro? "Referencia" (expresión que usé al comienzo y que usaré de ahora en adelante) es el nombre

4. Haack 1982: p. 77.
5. Para un análisis detallado de la relación en cuestión así como del interés de cierta rama de la lógica (la denominada "lógica informal") por la argumentación informal misma, véase Comesaña 1998, incluido en la Enciclopedia Lógica.

genérico que esta relación recibe cuando el lenguaje involucrado es el lenguaje natural (o mental).[6] La semántica filosófica es, por tanto, un capítulo tanto de la filosofía de la lógica como de la filosofía del lenguaje.

Ahora bien, ¿cuál es el panorama teórico desarrollado en torno a la noción de referencia? En otras palabras, ¿cuáles son las concepciones de la referencia a las que puede apelar un teórico contemporáneo para explicar el significado? Ante todo, cabe señalar que la reflexión sobre el significado, aun cuando no constituyese, como en nuestros días, una disciplina filosófica independiente -la semántica filosófica- ha apelado siempre al concepto de referencia. Dicho de otro modo, la explicación del significado ha involucrado típicamente una explicación de la relación entre los signos lingüísticos, por un lado, y los objetos del mundo, por el otro. El rechazo de la referencia es, como se verá más adelante, una cuestión contemporánea. Ahora bien, tal explicación ha sido dada en términos muy distintos. Veamos, muy someramente, cuáles han sido éstos.

Los filósofos modernos consideraron que la relación de referencia entre las palabras y los objetos no es directa sino que está mediada por entidades mentales, las ideas. Según los llamados "realistas representativos" -entre los que se cuentan fundamentalmente Descartes, Locke y Hume-, toda palabra significa una idea, y ésta, a su vez, es causada por un objeto semejante a ella. Las nociones de causalidad y semejanza son los conceptos utilizados para explicar la relación

6. Cabe aclarar que, en este trabajo, el concepto de *referencia* (o *designación*) es considerado un concepto *genérico* (relación entre términos en general y objetos) que comprende distintos conceptos *específicos*, tales como los de *denotación* (relación entre nombres propios o descripciones y objetos), *aplicación* (relación entre términos generales no relacionales y objetos), *cumplimiento* (relación entre términos generales relacionales y objetos). Es preciso aclarar, sin embargo, que no hay acuerdo al respecto (Donnellan, por ejemplo, considera que las descripciones no siempre son usadas para hacer *referencia* a individuos; el concepto de referencia no es por tanto para él un concepto genérico que sirva para caracterizar la dimensión semántica de todas las expresiones del lenguaje, como se pretende aquí; véanse Donnellan 1966, 1968, 1972). La diferencia de uso será debidamente aclarada toda vez que sea pertinente.

epistémica entre objetos e ideas: los objetos causan ideas en la mente semejantes a ellos, pero, a diferencia de ellos, accesibles directamente a la conciencia y constitutivas, por tanto, de los significados lingüísticos. El rol semántico de las ideas descansa entonces en su rol epistémico. Vale la pena destacar que, para estos filósofos, las únicas portadoras de significado (es decir, las únicas entidades de las que puede decirse que tienen significado) son las palabras del lenguaje natural: las ideas no tienen significado sino que son los significados de las palabras.[7]

Siglos más tarde, John Stuart Mill sostuvo que cierto tipo de expresiones, los nombres propios, se relacionan directamente con los objetos del mundo, sin la mediación de ideas. De acuerdo con esto, un nombre propio no significa una idea sino el objeto mismo designado por el nombre. Pero Mill no ha explicado en qué consiste la relación de designación o referencia entre nombres y objetos.

La propuesta de Gottlob Frege, en cambio, vuelve a introducir la existencia de una instancia mediadora entre unos y otros: esta vez no se trata de las ideas modernas sino de los sentidos, entidades abstractas, accesibles a todos los hablantes competentes de un lenguaje natural. El sentido de un nombre es definido, a grandes rasgos, en términos del modo en que el objeto referido se presenta a la mente y se identifica con el sentido de una descripción definida asociada con el nombre en cuestión por los hablantes competentes. La propuesta fregueana da origen, de este modo, a las denominadas "teorías descripcionales de la referencia".

Por último, en este siglo, muchos filósofos -tales como, por ejemplo, Saúl Kripke, Hilary Putnam y Keith Donnellan- rechazan fuertemente la propuesta fregueana y sus refinamientos. Nuevamente, se apela a la noción física de causalidad, pero esta vez para dar cuenta no de la relación epistémica entre ideas y objetos sino de la relación semántica entre palabras y objetos. Este cambio está sin duda basado

7. Por el contrario, en las teorías psicosemánticas contemporáneas, tales como las de Dretske y Fodor, las representaciones mentales, equiparables a las ideas modernas, son las portadoras de significado por excelencia -las palabras o representaciones lingüísticas sólo tienen significado en forma derivada.

en el desplazamiento contemporáneo del interés teórico por el conocimiento al interés por el lenguaje, característico del denominado "giro lingüístico". En otros términos, por un lado, a diferencia de Descartes y los empiristas británicos, estos filósofos intentan independizar la explicación del significado de las expresiones de la explicación de su rol epistémico; por otro lado, al igual que aquéllos, apelan en su teoría a la noción de causalidad. Sobre esta base, extienden la explicación milleana del significado de los nombres propios, según la cual el nombre refiere directamente a un objeto, a (por lo menos) una gran parte de los términos del lenguaje. Esta es la idea central de las actuales teorías causales de la referencia.

A su vez, existen tres versiones principales del enfoque causal: la histórica, la indicativa o informacional y la teleológica. Las primeras teorías aparecidas fueron históricas, tales como las propuestas por Kripke y Putnam, y fueron introducidas para explicar el significado de las palabras del lenguaje natural. De acuerdo con ellas, el referente de un nombre propio o de un término general está constituido por su causa histórica, a saber, un individuo particular (o una clase particular de individuos) con el cual la expresión está relacionada mediante una cadena histórico-causal; esta cadena conduce al llamado "bautismo inicial", es decir, el momento hipotético en el cual la palabra es fundada o anclada en el objeto en cuestión. Las teorías indicativas o informacionales, tales como las propuestas por Dennis Stampe y Fred Dretske, desarrolladas para el lenguaje del pensamiento, sostienen en cambio que un concepto refiere a su causa confiable, la cual es identificada sobre la base de nuestras capacidades discriminatorias naturales. Finalmente, según las teorías teleológicas, como las de Ruth Millikan y David Papineau, propuestas tanto para el lenguaje natural como para el pensamiento, la referencia es determinada en virtud de la causa final o télos de las representaciones, es decir, sus funciones biológicas o aquello para lo cual han sido naturalmente seleccionadas.

Cabe destacar que el enfoque causal, en cualquiera de sus versiones, en tanto se basa en el uso de una noción física como la de causalidad, parece ofrecer una respuesta científicamente aceptable al tradicional problema de la intencionalidad; en otros términos, la

noción clave de causalidad parece abrir el camino hacia una explicación de lo intencional en términos de lo físico, y de esta manera, romper el engorroso círculo constituido por las usualmente interdefinidas nociones intencionales -tales como las de significado, intención, creencia y demás estados mentales. De este modo, la semántica tiene la opción concreta de transformarse en una disciplina empírica, compatible con la visión fisicalista o naturalista del mundo -es decir, aquélla según la cual todos los hechos, incluidos los hechos característicamente intencionales, son, en última instancia, reductibles a hechos físicos. Sin embargo, como podrá apreciarse, la causalidad no resulta ser un concepto lo suficientemente poderoso como para cumplir el esperado rol: el enfoque causal, en sus distintas versiones, presenta problemas que sugieren claramente que la explicación de la referencia no puede depender exclusivamente de la noción de causalidad.

En síntesis, puede decirse que existen dos concepciones principales de la referencia, que sirven para caracterizar dos períodos distintos de la semántica contemporánea. Estas concepciones expresan los dos modos diferentes en los que se ha pensado que un signo, término o representación puede referir a un objeto: o bien en virtud de estar asociada con una descripción que selecciona al objeto referido -tradición fregueana o concepción descripcional- o bien en virtud de mantener una apropiada relación causal con aquél -tradición kripkeana o concepción causal.

A la luz de las consideraciones precedentes, resulta conveniente retomar la tesis principal del libro, enunciada al comienzo: la explicación del significado, tanto lingüístico como mental, requiere la integración de factores provenientes de distintas concepciones acerca de la referencia; en otras palabras, sólo una teoría mixta o ecléctica de la referencia puede dar cuenta de la variedad y riqueza del significado. Más específicamente, en este trabajo, me propongo defender la tesis de que tal teoría resulta de la combinación de una teoría teleológico-causal para algunos conceptos básicos del pensamiento y una teoría histórica, descriptivo-causal, para la mayor parte de los términos no básicos del lenguaje natural -y para algunos conceptos igualmente no básicos del pensamiento. La primera teoría, esto es,

la semántica primaria, es una teoría psicosemántica, mientras que la segunda es casi exclusivamente una teoría linguosemántica. Cabe aclarar que, en mi opinión, la construcción de la teoría en cuestión tiene un carácter programático: no se trata de una tarea acabada sino de un programa a desarrollar. Como se verá, hay ciertos problemas que exigen un alto grado de refinamiento de la teoría semántica, que no es en absoluto fácil de alcanzar.

Ahora bien, lo que sigue puede sin duda considerarse una implícita defensa de la denominada "concepción correspondentista de la verdad", es decir, la concepción tradicional según la cual la verdad consiste en la relación de correspondencia o adecuación entre el lenguaje (o el pensamiento), por un lado, y el mundo, por otro.[8] Como es sabido, esta concepción dio lugar a la teoría clásica o tradicional de la verdad como correspondencia, tal como se encuentra en las obras de Platón y Aristóteles.[9] La noción clásica o tradicional de correspondencia incluye las siguientes notas: (i) la idea de fundamentación del lenguaje en el mundo: la verdad exige que el lenguaje se adecue al mundo y no a la inversa; la correspondencia es, por tanto, una relación asimétrica, que sólo va del mundo al lenguaje; (ii) la idea de que partes específicas del lenguaje, como, por ejemplo, las oraciones, se relacionan con partes específicas del mundo, esto es, los hechos.[10] Para tomar un ejemplo conocido, se considera

8. En este punto, se opone claramente a Barrio 1998, incluido también en la Enciclopedia Lógica, por cuanto este último contiene un ataque a la concepción correspondentista -y una defensa, en cambio, de la concepción deflacionaria de la verdad.

9. Véanse, por ejemplo, Platón, *Sofista* 263b y Aristóteles, *Metafísica,* Libro IV, apartado VII: p. 125.

10. En este trabajo, no voy a profundizar en el complejo problema de los llamados "portadores de verdad", es decir, el problema de establecer qué tipo de entidades son aquéllas de las que se predica la verdad. Baste decir que puede tratarse de entidades lingüísticas, como las oraciones (tipo) y las emisiones (u oraciones-caso), de entidades abstractas, como las proposiciones o pensamientos (para algunos), o de entidades mentales, como las proposiciones (para otros) y las creencias. Presupondré entonces, por lo general, que los portadores de verdad son las emisiones u oraciones (caso), pero con la aclaración de que éstas pueden pertenecer tanto al

que el hecho de que la nieve es blanca, y sólo ése, es lo que determina la verdad de la oración "La nieve es blanca", y no a la inversa. A estos dos rasgos se ha agregado tradicionalmente un tercero, a saber, el carácter objetivo e independiente de lo mental del mundo con el que nos relacionamos por medio del lenguaje -lo cual hace del realismo metafísico parte constitutiva de las teorías tradicionales de la correspondencia.[11] De este modo, y dejando de lado el último punto, la concepción correspondentista tradicional puede caracterizarse como aquélla según la cual la verdad es definida como la relación asimétrica de adecuación entre oraciones específicas del lenguaje y hechos específicos del mundo. Puede decirse entonces que, desde el punto de vista correspondentista, la verdad es una propiedad real o sustantiva de las oraciones. De ahí que cualquier teoría que ejemplifique a esta concepción sea considerada una teoría explicativa de la verdad, es decir, una teoría que ofrece una explicación de la naturaleza de la verdad. Además, dado que la naturaleza en cuestión es la adecuación entre el lenguaje y el mundo y tal relación constituye una manera paradigmática de concebir a las propiedades semánticas, toda teoría correspondentista de la verdad es también considerada una teoría explicativa del significado. En otros términos, toda teoría de la correspondencia constituye la parte central de una teoría del significado o de la representación puesto que identifica a la verdad con una teoría de la verdad que identifica a esta última con la aptitud que

lenguaje natural como al pensamiento o lenguaje mental -en los términos anteriores, pueden ser tanto entidades lingüísticas como mentales. La relación entre el lenguaje y el pensamiento será tratada con cierto detenimiento en la conclusión general.

11. Este último punto es sin duda muy discutible. Más específicamente, cabe preguntarse si el realismo metafísico, lejos de ser parte constitutiva de la teoría de la correspondencia, no es acaso una doctrina metafísica conceptualmente independiente, que ha sido sistemáticamente asociada con aquélla a lo largo de la historia de la filosofía. De acuerdo con este punto de vista, la unión entre realismo y correspondencia no sería conceptual sino meramente coyuntural. Para una defensa de esta tesis, que excede los contenidos aquí desarrollados, véanse Devitt 1991b y Barrio 1998.

tienen las oraciones del lenguaje para representar a los hechos del mundo.[12]

Cabe preguntarse de qué manera se insertan las explicaciones de la referencia antes mencionadas en la estructura de la concepción correspondentista; en otros términos, de qué manera precisa se relaciona, en el marco de la correspondencia, el concepto de referencia con el de verdad. En este punto, es oportuno recordar que la caracterización tanto de la relación de correspondencia o adecuación como de la noción de hecho presenta serios problemas. Por un lado, los intentos por explicar la adecuación suelen recaer en metáforas de poco valor explicativo, como que consiste en que el lenguaje "refleje" o sea una "pintura" del mundo. Por otro, es difícil decir qué es exactamente un hecho. En primer lugar, los hechos constituyen entidades de estructura compleja, presuntamente isomorfas a las oraciones; y, en segundo lugar (y, justamente, como consecuencia de lo anterior), no parecen poder ser identificados sin utilizar el concepto de oración verdadera, que es lo que se intenta explicar.[13] Por

12. Cabe aclarar, sin embargo, que la identificación de la verdad correspondentista con el significado no fue explícita hasta que la reflexión sobre el lenguaje comenzó a ser prioritaria por sobre otras cuestiones filosóficas y adquirió la categoría de una disciplina filosófica independiente, esto es, en la segunda mitad del siglo pasado y, fundamentalmente, en este siglo. Fue Gottlob Frege quien utilizó por primera vez la noción de *valor de verdad* en la explicación del significado de las oraciones informativas (véase Frege 1892); Donald Davidson, por su parte, fue quien por primera vez identificó el significado de tales oraciones con sus *condiciones veritativas tarskianas* (véase Davidson 1967).
13. A modo de ejemplo, desde el punto de vista de la estrategia de los hechos, la oración (o) "El invierno ruso causó la primera gran derrota de Napoleón" es verdadera en virtud de su correspondencia con el hecho de que el invierno ruso causó la primera gran derrota de Napoleón. Pero la única manera de identificar el hecho en cuestión es mediante el uso de (o), es decir, especificando que se trata del hecho que se corresponde con (o) cuando ésta es *verdadera*. Por consiguiente, la explicación de la verdad que apela a la noción de hecho es circular: explica la verdad en términos de una noción, la de hecho, cuya explicación presupone la noción de verdad. Conectado con lo anterior, está el argumento conocido como *"sling shot"*. Véanse Davidson 1969 y 1967: p. 19, también está en Quine 1953: p. 161.

estas razones, gran parte de las versiones contemporáneas de la concepción correspondentista, versiones que llamaré "atomísticas" u "ortodoxas", entre las que se cuenta la aquí defendida, se suelen formular en términos como los siguientes:

Una oración es verdadera si y sólo si lo es en virtud de (i) su estructura sintáctica, (ii) las relaciones referenciales entre las palabras que la componen y los objetos del mundo y (iii) la naturaleza del mundo.[14]

La clave aquí es (ii). Por un lado, la vaga relación de adecuación entre el lenguaje y el mundo es explicada en términos de la relación de referencia entre las palabras y los objetos, la cual se considera que puede ser a su vez (por lo menos parcialmente) explicada en términos causales, ya sea histórico, indicativo o teleológico-causales -es decir, en términos científicamente aceptables. (Cabe destacar que la relación causal, en la medida en que es una relación asimétrica que va de los objetos a las palabras, recoge perfectamente la tradicional idea de fundamentación.) Por otro lado, la oscura noción de hecho es eliminada en favor de los objetos comunes y corrientes, que constituyen el extremo ontológico de las relaciones referenciales. Desde esta perspectiva, en términos del ejemplo antes mencionado, "La nieve es blanca" es verdadera en virtud de que (i) tiene la estructura sujeto-predicado, (ii) las palabras que constituyen el sujeto y el predicado mantienen ciertas relaciones referenciales con ciertos objetos del mundo o, en otros términos, hay un objeto denotado por el sujeto que pertenece a la clase de objetos a la que se aplica el predicado y (iii) la nieve es, en efecto, blanca. De este modo, se considera que la referencia, en la medida en que permite relacionar el lenguaje con entidades no lingüísticas con condiciones de individualización simples e independientes del concepto de verdad, es el concepto que otorga conte-

14. Si se acepta que el realismo es parte constitutiva de la teoría de la correspondencia, tal naturaleza debe concebirse como objetiva e independiente de lo mental, pues tales rasgos son los que caracterizan a la concepción realista del mundo.

nido preciso a la tradicional idea de adecuación o correspondencia. Es por esta razón que, como mencioné más arriba, los capítulos que siguen pueden ser considerados una implícita defensa de la concepción correspondentista de la verdad: la idea subyacente es que si se propone una adecuada explicación de la referencia, el proyecto de explicar la verdad en términos de la correspondencia entre el lenguaje y el mundo resulta viable.[15]

Antes de terminar esta introducción, quisiera entonces destacar dos aspectos del enfoque defendido que encuentro significativos. En primer lugar, el enfoque en cuestión forma parte de una perspectiva fáctica de la semántica, en tanto presupone una concepción de ésta última como disciplina que se ocupa de los signos (nuevamente, tanto del lenguaje natural como del pensamiento o lenguaje mental) y sus propiedades, a saber, paradigmáticamente, la referencia y la verdad. En este punto, se opone a la que puede denominarse "perspectiva interpretacionista", según la cual la semántica ha de ocuparse no de los signos sino de un tipo especial de conducta, característicamente humana, a saber, la conducta verbal o lingüística.[16] Como es sabido,

15. En términos más estrictos, según el punto de vista defendido, la teoría del significado tiene una estructura básica tarskiana, a la que se le agrega la explicación de la referencia. Para decirlo de otro modo, el punto de vista defendido es aquél según el cual la teoría del significado es el producto de la combinación de una teoría de la verdad semejante a la propuesta por Alfred Tarski para los lenguajes formales y una teoría de la referencia para el lenguaje natural y el lenguaje mental. La teoría aquí propuesta se ubica entonces en el conjunto de aquellas teorías referencialistas del significado que convierten a la noción tarskiana de satisfacción en el concepto (empírico) de referencia, definido fundamentalmente en términos causales. Para la definición tarskiana de la verdad en términos del concepto de satisfacción, véanse Tarski 1935 y 1944; para una defensa ya clásica del enfoque atomístico, esto es, la mencionada combinación de la estructura tarskiana y la explicación de la referencia en términos causales, véase Field 1972. Este tipo de complementación de la teoría tarskiana parece adecuarse perfectamente al ideal fisicalista del propio Tarski, de acuerdo con el cual la semántica debe formar parte de la visión fisicalista y, por tanto, científicamente aceptable, del mundo.

16. Cabe aclarar que los adjetivos "fáctico" e "interpretacionista" para calificar a los enfoques referencialista y conductista de la semántica respectivamente han sido tomados de Devitt 1991b: cap. 10.

esta última perspectiva está representada por las teorías semánticas de Willard Quine y Donald Davidson.[17] Considérese, por ejemplo, la teoría de la interpretación radical propuesta por éste último: de acuerdo con ella, el lenguaje es concebido no como un conjunto de signos con una dimensión semántica sino fundamentalmente como un conjunto de actos o conductas de emisión e interpretación de oraciones. En opinión de Davidson, a la luz de la evidencia de la conducta lingüística disponible al intérprete de un lenguaje natural, la referencia de los términos es inescrutable y, por esa razón, si el objetivo es proponer una teoría semántica empírica, el concepto de referencia debe ser abandonado.[18] De este modo, el proyecto de construir una teoría correspondentista de la verdad sobre la base de una explicación de la referencia es, para él, inviable. Sin embargo, Davidson considera que hay otra vía abierta: la teoría de la verdad puede dar lugar no ya estrictamente a una teoría del significado -puesto que no hay entidades que constituyan los significados, independientemente de nuestros actos de comunicarnos unos con otros- sino a una teoría de la interpretación lingüística que prescinde completamente de la noción de referencia. De ahí la propiedad del uso del adjetivo "interpretacionista" para calificar a esta perspectiva.

A la luz de lo anterior, resultará más clara la segunda característica del punto de vista aquí defendido que me interesa destacar: su carácter atomístico, en oposición al carácter holístico de la perspectiva interpretacionista. El enfoque atomístico, también denominado "building-block", puede ser caracterizado como aquél según el cual el punto de contacto entre la teoría semántica y la evidencia está situado en las asignaciones de valor semántico a las palabras o átomos del lenguaje; el enfoque holístico, por su parte, sitúa ese punto exclusivamente en las asignaciones de condiciones veritativas a las oraciones, emisiones, oraciones o todos oracionales -de ahí el uso del adjetivo "holístico". Es claro entonces que el giro hacia el

17. Véanse, por ejemplo, Quine 1960 y Davidson 1967.
18. La tesis de la inescrutabilidad de la referencia sirve de base al famoso argumento quineano de la indeterminación de la traducción. Véase Quine 1960: cap. 2.

holismo se basa en la antes mencionada eliminación de la referencia como concepto explicativamente útil en la teoría semántica. Si se considera, a la manera de Davidson, que no hay relaciones objetivas y empíricas entre las palabras y los objetos, entonces será lógico situar el único punto de contacto entre el lenguaje y el mundo -y, por tanto, entre la teoría y la evidencia- en el nivel de los todos oracionales.[19]

Ahora bien, en este punto, considero pertinente señalar muy brevemente cuáles son mis razones principales para adscribir a la perspectiva fáctica y atomística, y rechazar en cambio la perspectiva interpretacionista y holística. Cabe aclarar, sin embargo, que la fundamentación rigurosa de la postura adoptada excede ampliamente los límites de este trabajo. Fundamentalmente, considero que la perspectiva interpretacionista en general se basa en una concepción conductista del lenguaje, según la cual éste se reduce a un conjunto de conductas públicamente observables de asentir y disentir a oraciones frente a ciertas circunstancias del mundo, que no está justificada. El compromiso ontológico con el conductismo es, como vimos, un rasgo esencial de la perspectiva interpretacionista, y es precisamente este tipo de compromiso ontológico lo que determina su peculiar concepción de la evidencia disponible y el consiguiente rechazo de la referencia como noción empírica. Más específicamente, desde esta perspectiva, la referencia no puede ser definida en términos de una relación física (como la causalidad) porque la observación

19. Cabe destacar, además, que el punto de vista defendido da lugar, como sugerí anteriormente, a lo que puede denominarse "una versión *ortodoxa* de la teoría de la verdad como correspondencia". Por el contrario, el enfoque holístico constituye una versión *heterodoxa* de la teoría correspondentista, a saber, una versión que no apela a la noción de referencia sino a la de interpretación, introduciendo de este modo, junto con la perspectiva del intérprete, un conjunto de restricciones epistémicas en la teoría semántica. A fin de fundamentar este punto, es útil recordar que, para Davidson, la teoría de la interpretación logra aclarar la noción de verdad correspondentista de la única manera en que, según él, ésta puede ser aclarada, a saber, mostrando de qué manera es *usada* por los hablantes competentes de un lenguaje natural. Es, por tanto, al igual que toda teoría de la correspondencia, una teoría explicativa de la verdad. Véanse Davidson 1969 y 1977.

de la conducta lingüística antes mencionada, que da lugar al único tipo de evidencia considerado pertinente en función de la metodología semántica adecuada al compromiso conductista, no permite seleccionar una entre las múltiples relaciones físicas posibles entre el acto lingüístico y el entorno. Pero, a menos que se dé algún argumento independiente en favor de la concepción conductista del lenguaje, no hay razón para creer que la única evidencia que debe ser tomada en cuenta en la construcción de una teoría semántica sea evidencia de la conducta lingüística, en particular, de la actitud de asentimiento o disentimiento frente a oraciones. Y, si esto es así, no hay razón para creer que la referencia sea inescrutable en términos absolutos, es decir, en relación con toda evidencia posible. Cabe aclarar que el aquí sugerido abandono del marco conductista davidsoneano da lugar a una ampliación de la evidencia disponible -de modo tal de incluir no sólo evidencia de la actitud de asentir y disentir frente a determinadas oraciones sino también, por ejemplo, evidencia de la adscripción de propiedades semánticas a los términos de un lenguaje dado. Es claro que la tarea de explicar tales adscripciones en términos empíricos -más específicamente, en términos por lo menos parcialmente causales- puede presentar varias dificultades, entre las cuales cabe mencionar al problema de la ambigüedad, es decir, la existencia de múltiples correlatos ontológicos como posibles referentes de un término. Sin embargo, quienes piensan que la evidencia disponible no debe restringirse a aquélla basada en la observación de las conductas de asentir y disentir consideran que esta ambigüedad no es fatal -es decir, que hay distintos mecanismos que permiten eliminarla, tales como, por ejemplo, el uso de descripciones o la apelación a relaciones causales legales.[20] De este modo, considero que Davidson no ha ofrecido ninguna razón para que el enfoque atomístico de la correspondencia, en donde la referencia juega un rol central, deba ser abandonado.

20. Es interesante destacar nuevamente que este libro contiene una evaluación del problema de la ambigüedad muy distinta de la presentada en Barrio 1998, en donde se considera que se trata de un problema insoluble, que incita a abandonar toda expectativa correspondentista.

Por último, quisiera incluir una breve descripción de la estructura central de los capítulos que siguen, cada uno de los cuales consta de dos secciones principales. En la primera sección del primer capítulo, expongo las teorías descripcionales de la referencia (Frege 1892a, Russell 1905, Searle 1958, Strawson 1959), para luego destacar los problemas que determinaron el surgimiento de las teorías causales. La segunda sección está dedicada a la exposición de la primera versión de aquéllas, a saber, la teoría histórico-causal de Kripke (Kripke 1980), a continuación de lo cual examino los principales problemas que presenta. En el segundo capítulo me ocupo de las teorías mixtas surgidas a partir de la combinación de las anteriores, es decir, aquéllas que combinan factores descriptivos y factores histórico-causales. La primera sección está centrada en el análisis del primer tipo de teorías mixtas, a saber, las teorías del doble factor (Putnam 1975, Block 1986). La segunda sección contiene el examen de la teoría descriptivo-causal de Devitt y Sterelny (Devitt 1974, Devitt y Sterelny 1987, Devitt 1996). Por último, el tercer capítulo está dedicado a las teorías causales no históricas. En la primera sección, expongo y evalúo una teoría indicativo-causal, la teoría covariacional (Fodor 1987, 1990, 1994). En la segunda examino el enfoque teleológico-causal ejemplificado en la teoría de Millikan (Millikan 1984, 1995). Como conclusión, propongo, como señalé al comienzo, una teoría mixta, que reúne aspectos tanto de la teoría descriptivo-causal como de la teoría teleológica, en la línea de la propuesta de Sterelny (Sterelny 1990), pero distinta de esta última en algunos aspectos significativos.

Capítulo 1
La concepción descripcional y el modelo histórico-causal

El título de este capítulo alude a las dos concepciones principales de la referencia que se han ofrecido, las cuales sirven para caracterizar dos períodos distintos en la filosofía analítica del lenguaje. Estas concepciones expresan los dos modos diferentes en los que se ha pensado que una representación puede referir a un objeto: o bien en virtud de estar asociada con una descripción que selecciona a un único objeto, a saber, el referente -concepción descripcional o tradición fregueano-russelleana- o bien en virtud de mantener una apropiada relación causal con aquél -concepción causal o tradición kripkeana-. En lo que sigue, me ocupo entonces de ofrecer los lineamientos centrales de cada una de estas concepciones -más específicamente, de las teorías de Frege y Russell, por un lado, y de Kripke, por otro-, así como de presentar las críticas a cada una de ellas que encuentro pertinentes. Cabe aclarar que si bien las teorías presentadas -esto es, las teorías de Frege, Russell y Kripke- han sido ofrecidas fundamentalmente como teorías acerca del lenguaje natural, son perfectamente aplicables al pensamiento o lenguaje mental.[21] La

21. Cabe señalar que no hay acuerdo sobre este punto. A modo de ejemplo, Sterelny considera que en el caso de que la teoría de Kripke funcionase, debería hacerlo para el lenguaje del pensamiento (véase Sterelny 1990). Evans, en cambio, considera que sólo puede ser aplicada al lenguaje

conclusión obtenida es que no es posible sostener ni una teoría puramente descripcional, a la manera de Frege y Russell, ni una teoría puramente causal, a la manera de Kripke: en mi opinión, la explicación del mecanismo referencial, tanto del lenguaje como del pensamiento, exige, por el contrario, la combinación de factores descriptivos y factores causales. De ello me ocuparé, sin embargo, en el capítulo siguiente, es decir, después de evaluar y rechazar a las teorías puras de ambos signos.

1. La concepción descripcional

1.1. Los nombres propios: Frege vs. Mill

No es mi propósito hacer aquí una exposición detallada de la teoría de Frege sino tan sólo destacar sus tesis principales concernientes al significado de los nombres propios así como enfatizar la gran influencia que tuvieron en la filosofía del lenguaje contemporánea. Cabe aclarar que las razones de esta restricción son de índole puramente expositiva: dado que las tesis acerca de los otros tipos de términos son similares a sus tesis acerca de los nombres propios, he decidido tomar como ejemplo de la teoría a éstas últimas. Por lo demás, ése es el punto en el que la teoría se opone claramente a la de Mill, oposición que resulta sumamente iluminadora. Finalmente, es su concepción de los nombres la que es retomada y, en cierto sentido, perfeccionada por la teoría de las descripciones de Russell, de la que me ocupo en la sección siguiente.

Ahora bien, de acuerdo con lo anterior, la teoría de Frege acerca de los nombres propios puede ser mejor apreciada si se la compara con la teoría predominante en la época en que fue presentada, a saber, la teoría de Mill.[22] Según este último, los nombres propios son expresiones puramente denotativas, es decir, su significado está

natural (véase Evans 1982). Las razones en favor de este último punto de vista me resultan sumamente oscuras.

22. Véase Mill 1867: Libro I, cap. II.

constituido exclusivamente por el individuo referido en cada caso, esto es, en los términos más precisos de la teoría, por su denotado. A modo de ejemplo, el significado de "Tolstoi" es Tolstoi, es decir, la persona referida por el nombre en cuestión. Un nombre funciona, respecto de un individuo, como una etiqueta. Por consiguiente, se trata de una teoría que, en términos contemporáneos, puede ser denominada "teoría de la referencia directa": no hay ninguna instancia intermedia entre el nombre y el individuo referido; el primero refiere directamente al segundo.

Frege, por el contrario, sostiene que el significado de un nombre propio involucra dos dimensiones distintas. Por un lado, está el individuo referido o denotado por el nombre, dimensión que constituye el referente o denotado -también llamado "referencia" o "denotación".[23] Por otro lado, está el modo en que el individuo en cuestión se presenta a la mente, modo que está constituido por la descripción de cierta propiedad identificadora del individuo; esta segunda dimensión constituye el sentido del nombre. En términos más estrictamente fregueanos, el sentido de un nombre propio está constituido por el sentido de una descripción definida -que selecciona una propiedad identificadora del individuo referido- asociada con el nombre en cuestión. Los nombres tienen no sólo denotado o referente sino también sentido. Más exactamente, en palabras del propio Frege:

Un nombre propio (una palabra, un signo, una combinación de signos, una expresión) expresa su sentido y denota o designa su

23. Cabe aclarar que, en Frege, a diferencia de lo que señalado en la introducción, el concepto de denotación es equivalente al concepto de referencia. La razón de ello es que el hecho de que Frege concibe el funcionamiento semántico de todas las expresiones bajo el modelo de los nombres propios; en otros términos, todas las expresiones funcionan semánticamente como nombres de (algún tipo de) objetos. De este modo, según Frege, los términos generales nombran *conceptos* por lo que éstos últimos son el tipo de objeto denotado por los términos generales; las oraciones, como vimos, denotan un tipo peculiar de objeto, constituido por sólo dos unidades, *lo Verdadero* y *lo Falso*. Véase Frege 1892b.

denotación. Por medio de un signo, expresamos su sentido y designamos su denotación.[24]

Retomando el ejemplo anterior, el significado de "Tolstoi" está constituido en parte por Tolstoi, la persona referida, y en parte por el modo de presentación de la misma, especificado por la descripción identificadora "El autor de *Ana Karenina*", asociada con el nombre por los hablantes competentes.

La noción de sentido cumple entonces un rol clave en la teoría fregueana: a diferencia de Mill, para quien los nombres propios son expresiones puramente denotativas, es decir, expresiones cuyo significado es agotado por el objeto denotado o referido, Frege les reconoce una dimensión connotativa: en la medida en que los nombres tienen sentido, connotan cierta propiedad o característica del objeto referido. De este modo, la concepción fregueana de los nombres es semejante a la concepción milleana de los términos generales, según la cual éstos últimos son expresiones no sólo denotativas sino también connotativas.

En lo que sigue, intentaré entonces profundizar en la noción de sentido, tal como es presentada por Frege, quien la caracteriza fundamentalmente de tres maneras distintas:

Es natural pensar que con un signo (un nombre, una combinación de palabras, un grafismo) está conectado, además de lo designado por él, lo que puede llamarse la denotación del signo, lo que yo denominaría el sentido del signo, en el cual está contenido el modo de presentación.

El sentido de un nombre propio es aprehendido por todo aquél que tiene suficiente familiaridad con el lenguaje o con la totalidad de las designaciones de las que el nombre propio es parte [...].
De esta manera, la imagen se distingue esencialmente del sentido del signo, que puede ser propiedad común de muchos y que, en

24. Simpson (comp.) 1973: p. 9.

consecuencia, no es parte o modo de la mente individual. Es difícil negar que la humanidad posee un tesoro común de pensamientos que son transmitidos de una generación a otra.[25]

Ante todo, cabe destacar que entre estas caracterizaciones existe una suerte de tensión. Las dos primeras, al describir al sentido como el modo de darse del objeto -o, más exactamente, como aquello que contiene al modo de darse- y como aquello que comprende el hablante competente, parecen aludir a una entidad subjetiva, individual y psicológica. Según la tercera, en cambio, se trata de algo común, no sólo a todos los miembros de una cierta comunidad lingüística sino a la humanidad en su conjunto. Frege, cuyo rival histórico es el psicologismo -encarnado también en la figura de Mill-, se preocupa por enfatizar esta última caracterización: uno de sus objetivos principales al presentar una teoría semántica es legitimar el carácter objetivo del lenguaje y el pensamiento, rasgos que a su vez se relacionan con la presencia en ellos de una estructura lógica. Por esta razón, le interesa oponer la noción de sentido, central en su teoría, a la de imagen o representación asociada con un nombre: mientras que éstas últimas son subjetivas, individuales (en el sentido de que varían según el individuo) y de naturaleza mental o psicológica, los sentidos son objetivos (o, más precisamente, intersubjetivos), comunes a todos los seres humanos y de naturaleza abstracta.

Una manera de ser fiel al espíritu del texto de Frege y eludir al mismo tiempo el compromiso con la existencia de entidades abstractas, a la manera de Platón, consiste en concebir a los sentidos como representaciones conceptuales -esto es, entidades mentales o psicológicas- pero cuidando de evitar su identificación con las imágenes. En esta línea, Salmon distingue las siguientes tres dimensiones atribuibles al sentido fregueano:

(i) el sentido como representación conceptual o concepto, y, por tanto, como entidad psicológica;

25. Las citas corresponden a Simpson (comp.) 1973: pp. 4-5, p. 5 y p. 7, respectivamente.

(ii) el sentido como lo que determina la referencia, y, por tanto, como unidad semántica;

(iii) el sentido como lo que determina la contribución de un término al valor informativo de la oración en la que figura, y por tanto, como unidad epistémica.[26]

De acuerdo con esto, el sentido de un nombre propio puede ser concebido como un concepto descriptivo (o expresable mediante una descripción) asociado con el nombre por los hablantes competentes, concepto que se distingue del concepto constituido por la imagen en virtud de sus funciones semántica y epistémica.[27]

La manera entonces de distinguir al sentido de la imagen es tomar en cuenta las antes mencionadas funciones semántica y epistémica del primero. Centrémonos por el momento en la primera de ellas: la función semántica del sentido es determinar el referente. Volviendo a nuestro ejemplo, el referente de "Tolstoi" está determinado

26. Véase Salmon 1982: Parte I, cap. 1.

27. Véase el siguiente comentario de Putnam sobre el mismo punto: "Si nuestra interpretación de la doctrina tradicional acerca de la intensión y la extensión hace justicia a Frege y Carnap, entonces toda la cuestión acerca del psicologismo y el platonismo resulta ser de alguna manera como ahogarse en un vaso de agua, en lo que respecta a la teoría semántica. (Por supuesto, es de suma importancia en lo que respecta a la filosofía de las matemáticas.) Puesto que incluso si los significados son, como lo consideran Frege y Carnap, entidades 'platónicas' en lugar de entidades 'mentales', 'captar' tales entidades es probablemente un estado psicológico (...) Más aún, el estado psicológico determina unívocamente a la entidad 'platónica'. Por tanto, parece ser de alguna manera un asunto convencional el tomar ya sea a la entidad 'platónica', ya sea al estado psicológico, como el 'significado'. Y considerar que el significado está constituido por el estado psicológico difícilmente tenga la consecuencia temida por Frege, a saber, que el significado deja de ser público. Puesto que los estados psicológicos son 'públicos' desde el momento en que personas distintas (y aun personas de distintas épocas) pueden estar en el *mismo* estado psicológico. Verdaderamente, *el argumento de Frege en contra del psicologismo es sólo un argumento en contra de la identificación de conceptos con particulares mentales, no con entidades mentales en general.*" Putnam 1975: p. 222. Las bastardillas son mías.

por el sentido de "El autor de *Ana Karenina*" o, en otros términos, por el concepto (descriptivo) EL AUTOR DE *ANA KARENINA*, por cuanto "Tolstoi" refiere a la persona que de hecho posee la propiedad seleccionada por la descripción, esto es, la persona que es el autor de Ana Karenina; en otras palabras, la manera de determinar el referente de "Tolstoi" es identificar a aquella (única) persona que posee la propiedad de ser el autor de *Ana Karenina*.[28] Desde este punto de vista, la relación de un nombre con su referente no es directa, como sostiene Mill, sino que está mediada por el sentido: el sentido oficia de intermediario entre el nombre y su referente; es por medio del sentido expresado que el nombre denota o refiere a un individuo. Por lo tanto, la función semántica descripta no sólo permite delimitar la noción de sentido sino que determina la peculiar concepción del mecanismo referencial que caracteriza a las teorías descripcionales. Como se mencionó en la introducción, en general, según la concepción descripcional de la referencia, el mecanismo en virtud del cual una representación refiere a un objeto es su asociación con una descripción que se aplica (únicamente) al objeto en cuestión.

Y éste es exactamente el punto en donde la teoría de Frege presenta una ventaja indiscutible por sobre la de Mill: mientras que la teoría de Mill no ofrece explicación alguna del mecanismo en virtud del cual un nombre refiere a un individuo -simplemente afirma que lo hace de una manera directa-, la de Frege sí lo hace. La teoría de Frege, a diferencia de la de Mill, contiene una explicación del mecanismo de referencia de los nombres, a saber, en términos de descripciones asociadas.

Además, cabe aclarar que su peculiar manera de concebir la relación entre el sentido y la referencia permite calificar a la teoría fregueana como una teoría representacionalista del significado de los nombres, es decir, como una teoría que considera que no hay propiedades semánticas que no contribuyan a la determinación de la

28. Cabe señalar que las mayúsculas son la convención elegida para designar conceptos o representaciones mentales.

referencia de los nombres. En otras palabras, si bien es posible afirmar que Frege se compromete con la existencia de una propiedad semántica, el sentido, que no es idéntica a la propiedad de referir a un objeto, se trata de una propiedad cuya función semántica es determinar el referente, a saber, la propiedad de referir a un objeto de un modo determinado o bajo un cierto modo de presentación del mismo. Por consiguiente, puede decirse que, para Frege, todas las propiedades semánticas son representacionales, es decir, fundamentan la aptitud representacional del lenguaje -en virtud de la cual éste representa o refleja al mundo.[29]

Ahora bien, la teoría fregueana del significado de los nombres propios, cuyas tesis principales se acaban de destacar, ha sido frecuentemente asociada con una teoría de la comprensión de los nombres en cuestión. Esta última, a diferencia de la primera, no se refiere a los significados como propiedades de los nombres sino al proceso de comprensión de los mismos por parte de los hablantes (e intérpretes) de un lenguaje. La tesis de que los nombres expresan un sentido y denotan un referente tiene su correlato en la tesis de que comprender un nombre implica captar su sentido y, consiguientemente, ser capaz de acceder al referente. Ahora bien, en la interpretación de Salmon, según la cual, los sentidos son conceptos (de descripciones identificadoras), la teoría del significado parece identificarse con la teoría de la comprensión, puesto que poseer un concepto es o equivale a comprenderlo. ¿Qué puede querer decir que uno posee un concepto que no comprende? Una teoría sobre la posesión de conceptos es una teoría sobre la comprensión de conceptos; del mismo modo, bajo el supuesto de que los sentidos son conceptos, la explicación fregueana del significado en términos de la posesión de un sentido involucra una teoría de la comprensión del sentido.

29. Es preciso aclarar que, estrictamente, Frege no habla de *propiedades* semánticas sino de *entidades* -el sentido y la referencia- asociadas con los distintos tipos de expresiones (véase, por ejemplo, lo que se dice más adelante acerca de las oraciones y los valores de verdad). Pero la consideración de esta diferencia ontológica no modifica en absoluto el sentido de la tesis atribuida al autor en este párrafo.

En mi opinión, este último punto está relacionado con la otra función del sentido antes mencionada, a saber, su función epistémica, según la cual el sentido determina la contribución de un término al valor informativo de la oración en la que figura. De acuerdo con esto, el sentido involucra información, es decir, conocimiento. Veamos de qué manera lo hace en términos del ejemplo anterior. El significado del nombre "Tolstoi" involucra, para Frege, un sentido, a saber, el concepto descriptivo EL AUTOR DE *ANA KARENINA* o, en otros términos, el concepto de la descripción "El autor de *Ana Karenina*" asociada con aquél; esto involucra a su vez la creencia de que el individuo referido por el nombre tiene la propiedad seleccionada por la descripción, es decir, la creencia de que Tolstoi es el autor de Ana Karenina; ahora bien, dado que la creencia en cuestión es verdadera -de lo contrario, la descripción asociada no serviría para identificar al individuo referido-, no hay razón alguna para no considerarla conocimiento. Por consiguiente, la teoría fregueana del sentido implica que el significado de un nombre involucra conocimiento -en particular, el conocimiento de una propiedad identificadora- acerca del individuo referido. Es por esta razón que el sentido constituye no sólo una unidad semántica sino también una unidad epistémica. Este rasgo es muy significativo puesto que indica un alto grado de participación de factores epistémicos en la teoría semántica: se concede un rol epistémico a la unidad semántica fundamental de la teoría.

Por último, es preciso tener en cuenta, como señalé al comienzo, que la teoría presentada no se restringe al conjunto de los nombres propios sino que comprende también a las descripciones definidas, los predicados (términos generales contables, como "gato", y términos generales de masa, como "agua") y las oraciones -concebidas éstas últimas sobre el modelo de los nombres. De todos estos tipos de expresiones se dice, por tanto, no sólo que refieren a objetos (en el sentido general del término) sino que expresan un sentido; más específicamente, que refieren a objetos por medio de los sentidos expresados. En el caso de las oraciones, los sentidos expresados son pensamientos o proposiciones y los objetos referidos son los valores de verdad, concebidos como dos extrañas entidades: lo Verdadero y lo Falso. (Esto último

tiene la poco deseable consecuencia de que todas las oraciones verdaderas refieren al mismo objeto, así como todas las oraciones falsas.) Frege constituye entonces uno de los primeros filósofos que hace uso explícito del concepto de verdad para explicar el significado de las oraciones. Cabe recordar que Frege defiende el principio de composicionalidad, tanto para la dimensión del sentido como para la dimensión referencial: de este modo, el pensamiento expresado por una oración está determinado por los sentidos expresados por los componentes oracionales y el valor de verdad de la misma está a su vez determinado por la referencia de sus partes.

Hecha la aclaración anterior, es oportuno destacar que, mediante la introducción del concepto de sentido, Frege logra liberar a la semántica de un conjunto de problemas tradicionales, todos ellos derivados de la identificación del significado de un nombre con el individuo denotado o referido, a la manera de Mill. Es justamente por ser capaz de ofrecer una solución a estos problemas por lo que la teoría reemplazó a su antecesora milleana. A continuación, describiré brevemente algunos de los problemas en cuestión y de qué manera encuentran una solución en el marco de la teoría de Frege.

El problema de la identidad. Si los nombres propios significan aquello a lo que refieren, entonces los enunciados de identidad entre nombres resultan o bien falsos o bien, si son verdaderos, triviales; pero hay enunciados de identidad verdaderos que, lejos de ser triviales, tienen valor informativo: este hecho no puede entonces ser explicado en el marco de una teoría de la referencia directa. Un ejemplo puede ayudar a clarificar la idea anterior: si la referencia agota el significado de un nombre y "Charles Dodgson" y "Lewis Carroll" refieren al mismo individuo, no hay manera de explicar la diferencia de valor informativo entre la trivial

Charles Dodgson es Charles Dodgson

y la altamente informativa

Charles Dodgson es Lewis Carroll

Es el reconocimiento del sentido como parte esencial del significado de un nombre lo que hace posible explicar porqué el significado de "Charles Dodgson" es distinto del significado de "Lewis Carroll" y, consiguientemente, porqué el valor informativo de la primera de las oraciones mencionadas es distinto del de la segunda -como se recordará, según Frege, el sentido de un nombre es lo que contribuye a determinar el valor informativo de la oración en la que figura.

La paradoja de la denotación. La teoría de la referencia directa es incompatible con el principio de sustituibilidad de términos idénticos, según el cual dos términos que tienen el mismo significado pueden ser intercambiados en todos los contextos oracionales *salva veritate*, es decir, sin que cambie el valor veritativo de las oraciones resultantes. También este problema resulta más claro si se lo ilustra con un ejemplo, inspirado en el texto de Frege: "la estrella matutina" y "la estrella vespertina" tienen la misma referencia (Venus) y, por tanto, desde el punto de vista de una teoría de la referencia directa, el mismo significado; sin embargo, María, que ignora que la estrella vespertina es la estrella matutina, cree que la estrella vespertina protege la ciudad durante la noche pero no cree que la estrella matutina lo haga; luego,

> María cree que la estrella vespertina protege la ciudad durante la noche

es verdadera mientras que

> María cree que la estrella matutina protege la ciudad durante la noche

es falsa; en otros términos, los términos correferenciales "la estrella matutina" y "la estrella vespertina" no pueden ser sustituidos *salva veritate* en contextos como los contextos de creencias -conocidos como "opacos" u "oblicuos". Dada la generalidad del principio en cuestión, parecería que, por lo menos en algunos contextos, el

significado de los nombres propios no puede estar constituido por los individuos denotados o referidos. La noción de sentido se presenta entonces como una noción adecuada para desempeñar el rol en cuestión: en contextos oblicuos, los nombres adquieren un significado distinto, constituido básicamente por su sentido -en términos más estrictos, se trata de un significado constituido por un referente que es el sentido habitual y un sentido denominado "sentido indirecto". El uso de la noción de sentido, a diferencia de la asignación de meros referentes, permite entonces preservar el principio de sustituibilidad enunciado al comienzo.

El problema de las oraciones acerca de entidades ficticias. Introduciré este problema mediante un nuevo ejemplo: dados los presupuestos de la teoría de la referencia directa y la aceptación del mencionado principio de composicionalidad -según el cual el significado del todo oracional depende de los significados de sus componentes suboracionales-, no parece posible poder explicar la significatividad de oraciones tales como

Orfeo ha sido perdonado por los dioses

donde "Orfeo" refiere a Orfeo y Orfeo no existe, esto es, oraciones cuyos términos sujeto refieren a entidades ficticias y, por tanto, inexistentes, o, en otras palabras, oraciones cuyos términos sujeto no refieren a nada. La solución aportada por Frege reside en considerar a tales oraciones significativas pero carentes de valor de verdad, en la misma medida en que puede considerarse que sus términos sujeto (tales como "Orfeo"), si bien carecen de referencia, poseen, en cambio, sentido. Nuevamente, el concepto de sentido desempeña un rol clave en la solución del problema.

El problema de las oraciones existenciales negativas. En el planteo de este problema, haré uso del mismo ejemplo anterior: en el marco de la teoría de la referencia directa y dado el principio de composicionalidad, si "Orfeo" refiere a Orfeo y Orfeo no existe, no parece posible poder asignar valor de verdad alguno a

Orfeo no existe

En otras palabras, no parece posible afirmar que el valor de verdad -y, más específicamente, la verdad- de una oración como la anterior está determinada por la referencia de los componentes suboracionales, por cuanto su término sujeto no refiere a nada.

No es *prima facie* plausible aplicar a este tipo de oraciones la solución anterior (es decir, sostener que las oraciones existenciales negativas, al igual que las oraciones acerca de entidades ficticias, son significativas pero carentes de valor de verdad): dado que muchas de ellas son claramente verdaderas (como, por ejemplo, la antes mencionada "Orfeo no existe"), resulta antiintuitivo afirmar que todas ellas carecen de valor de verdad. Dentro del marco proporcionado por la teoría de Frege, se ha sugerido que las oraciones en cuestión deben ser interpretadas como oraciones acerca no de objetos sino de conceptos (esto es, sentidos), de los cuales se dice que representan conjuntos vacíos, es decir, conjuntos cuya extensión carece de miembros. De este modo, una oración como la anterior es considerada equivalente a

El conjunto representado por el concepto ORFEO no tiene ningún miembro en su extensión

de la cual puede decirse, sin inconvenientes, que es verdadera.[30]

En síntesis, las características de la teoría fregueana que me interesa destacar son las siguientes. En primer lugar, involucra una teoría del significado para los nombres propios y términos generales, según la cual unos y otros tienen, además de referentes, sentidos -entendidos en términos de los sentidos de las descripciones identificadoras asociadas en cada caso-, es decir, no es una teoría de la referencia directa. En segundo lugar, y como consecuencia de lo anterior, es una teoría descripcional de la referencia,

30. Véase Simpson 1964: cap. IV.

por cuanto sostiene que el mecanismo por el cual un término refiere a un objeto involucra el uso de una descripción asociada que selecciona una propiedad identificadora del mismo; cabe destacar que el factor descriptivo involucrado, en la medida en que sirve para caracterizar el conocimiento del hablante, introduce una dimensión epistémica en el concepto de significado. Por último, es una teoría representacionalista del significado, por cuanto sostiene que todas las propiedades semánticas de un término -esto es, las propiedades constitutivas del significado- contribuyen a determinar su referencia.

1.2. La teoría de las descripciones de Russell

La teoría de Russell es un típico ejemplo de teoría semántica atomística, pero no sólo en el sentido mencionado en la introducción, según el cual se trata de una teoría que contiene una explicación de la referencia, sino también en un nuevo sentido, distinto del anterior, según el cual es una teoría en la que el significado de la mayor parte de las expresiones del lenguaje es explicado en términos del significado de un subconjunto de átomos. Involucra, por consiguiente, tanto una especificación de cuáles son los átomos, como una teoría del significado para ellos y una teoría de la equivalencia semántica entre las expresiones atómicas y las no atómicas.

En cuanto a lo primero, la base de identificación de los átomos es gnoseológica: son expresiones atómicas aquéllas cuyos significados se conocen de manera directa (*by acquaintance*). Tenemos conocimiento directo tanto de datos sensoriales como de universales; luego, los términos que expresan ese tipo de conocimiento, a saber, los nombres propios y los términos generales, son atómicos.

En lo que concierne a la teoría semántica para los átomos, se trata de una teoría semejante a la teoría propuesta por Mill para los nombres propios. Para Russell, el significado tanto de los nombres propios como de los términos generales está constituido exclusivamente por los objetos por ellos referidos o denotados, datos sensoriales y

universales, respectivamente.[31] El significado de los átomos es entonces explicado en términos de la relación semántica de referencia directa -la cual se basa en la relación gnoseológica de conocimiento directo. En este punto, Russell se aleja entonces de la propuesta fregueana antes considerada.

Respecto del tercer punto, la tesis de equivalencia semántica entre las expresiones atómicas y no atómicas está contenida fundamentalmente en la denominada "teoría de las descripciones". Según Russell, a diferencia de los nombres propios y los términos generales, las descripciones definidas involucran un conocimiento no directo sino indirecto de los objetos referidos o denotados (*knowledge by description*). Por consiguiente, el significado de las descripciones no está constituido por los objetos por ellas denotados, en otras palabras, no es explicado en términos de la relación semántica de referencia directa: las descripciones son, por excelencia, las expresiones no atómicas del lenguaje, cuyo significado debe explicarse en términos del significado de las expresiones atómicas.

Veamos los detalles de la teoría. Según una exposición informal, una oración cuyo sujeto gramatical es una descripción definida tal como

(o) El autor de La Ilíada y la Odisea es el fundador de la épica clásica

es semánticamente equivalente a la conjunción de las siguientes oraciones:

(i) Existe *por lo menos* un individuo que es autor de *La Ilíada y la Odisea*

(ii) Existe *a lo sumo* un individuo que es autor de *La Ilíada y la Odisea*

(iii) Todo individuo que es autor de *La Ilíada y la Odisea* es el fundador de la épica clásica

31. Cabe destacar que para Russell, al igual que para Frege, los conceptos de referencia y denotación son sinónimos.

En una oración,

(o') Existe un individuo y sólo uno que es autor de La Ilíada y la Odisea y ese individuo es el fundador de la épica clásica

En el análisis, la descripción definida ("el autor de *La Ilíada y la Odisea*") ha desaparecido, dando lugar a una oración general existencial, en la cual ciertas propiedades (ser el único autor de *La Ilíada y la Odisea* y ser el fundador de la épica clásica) son predicadas, mediante el uso de términos generales ("autor de *La Ilíada y la Odisea*", "fundador de la épica clásica"), de un individuo indeterminado.

La tesis de Russell en apoyo de este análisis es que las oraciones del lenguaje natural como (o) ocultan su forma lógica bajo una engañosa apariencia gramatical: la estructura de sujeto y predicado no refleja la forma lógica, que es puesta de manifiesto mediante oraciones generales existenciales del tipo de (o'). Son éstas, reveladoras de la forma lógica, las únicas que permiten identificar los componentes semánticamente pertinentes de las oraciones. De este modo, la ausencia de descripciones definidas (y, en general, de frases denotativas) en oraciones del lenguaje lógico muestra que aquéllas no son componentes genuinos de las proposiciones o pensamientos en cuyas expresiones verbales aparecen.[32] En términos más simples, el análisis lógico de las oraciones con descripciones definidas pone de manifiesto el hecho de que no hay nada que

32. Es preciso señalar que la teoría russelliana de las descripciones forma parte de una teoría más general acerca de las denominadas "frases denotativas", de las que las descripciones definidas constituyen un subconjunto. En términos de Russell: "Entiendo por 'frase denotativa' una frase como cualquiera de las siguientes: un hombre, algún hombre, cualquier hombre, cada hombre, todos los hombres, el actual rey de Inglaterra, el actual rey de Francia, el centro de masa del sistema solar en el primer instante del siglo XX, la revolución de la Tierra alrededor del Sol, la revolución del Sol alrededor de la Tierra." Simpson (comp.) 1973: p. 29. Como señala Simpson, [...] en este contexto, 'frase denotativa' no debe interpretarse como 'frase que denota', sino como una expresión técnica cuyo significado se agota en la enumeración dada, sobreentendiéndose que en lugar de 'hombre' o 'rey de Inglaterra' pueden ir otros términos [...]". Simpson 1964: p. 75.

constituya el significado de esas expresiones -por lo tanto, aun cuando denoten, el objeto por ellas denotado no constituye su significado. Russell sostiene que se trata de expresiones que no tienen significado autónomo sino que contribuyen al significado de las oraciones en las que aparecen: podría decirse que su contribución se realiza mediante su desaparición y la consiguiente revelación de la forma lógica de las respectivas oraciones.

En síntesis, Russell extiende a las oraciones con descripciones definidas el análisis fregueano de las oraciones con frases denotativas en términos de oraciones cuantificadas: las oraciones que contienen descripciones definidas son analizadas en términos de oraciones generales existenciales que no contienen sino variables y términos generales. El significado de éstas últimas es especificado en términos de los significados de sus componentes, los únicos componentes genuinos de las oraciones en cuestión, a saber, los términos generales, los cuales, como vimos, están dados por los objetos por ellos denotados en forma directa -como vimos también, universales.

Ahora bien, es necesario destacar que para Russell, dado que sólo tenemos conocimiento directo de escorzos de objetos (datos sensoriales o *sense data*) pero no de objetos enteros (como, por ejemplo, las personas), los únicos nombres propios en sentido estricto (lógico, según Russell) son los demostrativos "esto" y "eso". En tanto no tenemos conocimiento directo de los objetos denotados por los nombres propios corrientes, su significado no puede estar constituido por aquello que denotan; si esto es así, los nombres propios corrientes se asemejan más a descripciones definidas que a verdaderos nombres. Para Russell, los nombres propios corrientes no son sino descripciones definidas abreviadas.

De este modo, por un lado, al igual que Frege, Russell considera que la explicación del significado de los nombres propios corrientes involucra la especificación del significado de una descripción asociada con el nombre, es decir, la especificación de ciertas propiedades que permiten identificar al individuo denotado o referido por el nombre. En otros términos, al igual que Frege, suscribe una teoría descripcional de la referencia de los nombres. Pero, por otro, a diferencia de Frege, considera que el significado de las descripciones debe explicarse en términos del significado de los términos generales, y el

de éstos últimos, a su vez, en términos de una relación de referencia directa con ciertos aspectos del mundo (universales). Estas tesis se basan en la suscripción, por parte de Russell, de ciertos presupuestos que Frege no comparte: por un lado, la extensión del análisis cuantificacional a las oraciones con descripciones definidas; por otro, la aceptación de la existencia de relaciones semánticas directas entre, por lo menos, (ciertos tipos de) expresiones y (ciertos tipos de) objetos, basadas a su vez en relaciones epistémicas directas. Para Russell, los nombres propios corrientes no involucran una relación directa con objetos del mundo, pero pueden descomponerse en descripciones que a su vez pueden descomponerse en términos generales que sí lo hacen. Para Frege, en cambio, ninguna instancia del lenguaje permite este acceso (semántico, más allá de que también involucre un acceso epistémico) directo al mundo.

Cabe preguntarse en este punto si esta teoría ofrece una solución a los antes mencionados problemas semánticos tradicionales. Me referiré brevemente y a modo de ejemplo a dos de ellos -la paradoja de la denotación y el problema de las oraciones existenciales negativas. En primer lugar, la paradoja de la denotación es solucionada mediante la introducción de la distinción entre figuración primaria y figuración secundaria de una descripción definida. Para tomar un ejemplo de Russell, la oración

Jorge IV quiso saber si Scott era el autor de Waverley

puede interpretarse o bien como

(I1) Existe un individuo y sólo uno que escribió Waverley y Jorge IV quiso saber si Scott era ese individuo

en donde la descripción definida tiene figuración primaria, es decir, no pertenece al contexto gobernado por el verbo de actitud proposicional; o bien como

(I2) Jorge IV quiso saber si sólo un individuo escribió Waverley y si ese individuo era Scott

en donde la descripción definida tiene figuración secundaria, por cuanto pertenece al contexto opaco u oblicuo, esto es, al contexto gobernado por el verbo de actitud proposicional. Ahora bien, Russell sostiene que en (I2), es decir, si la descripción definida tiene figuración secundaria, la sustitución de "el autor de Waverley" por "Scott" no puede hacerse *salva veritate* porque no se trata de términos de igual significado. Pero, en (I1), es decir, en el caso en que la descripción definida involucrada tiene figuración primaria, "el autor de Waverley" puede ser sustituida por "Scott" sin perjuicio del valor veritativo de la oración -piénsese en una situación en la que que el rey ve de lejos a alguien y pregunta si ese individuo, que es de hecho Scott, es Scott: sería un caso en que es verdad que el rey quiere saber si Scott es Scott.[33]

La teoría ofrece también una solución clara al problema de las oraciones existenciales negativas, con un mínimo compromiso ontológico -esto es, sin comprometerse con sentidos fregueanos ni, mucho menos, con la existencia de entidades ficticias correspondientes a sus respectivos nombres. Una oración como la antes mencionada

Orfeo no existe

es analizada en términos de otra que contiene una descripción definida tal como

El esposo de Eurídice no existe

y ésta última, es a su vez analizada, de acuerdo con la teoría de las descripciones, en términos de

No existe un individuo y sólo uno tal que ese individuo sea esposo de Eurídice.

33. Cabe destacar que la distinción entre figuración primaria y figuración secundaria de una descripción permite solucionar también un nuevo problema, denominado "paradoja del tercero excluido". Véase Russell 1905.

En síntesis, los puntos que me interesa destacar de la teoría de Russell son los siguientes. La teoría de las descripciones constituye fundamentalmente una teoría descripcional de la referencia de los nombres propios corrientes -que pertenecen al subconjunto de las expresiones no atómicas del lenguaje. En este sentido, constituye una extensión de la teoría fregueana de los nombres propios que vimos en el apartado anterior. Ahora bien, Russell agrega la mencionada tesis de equivalencia semántica, según la cual el significado de una oración singular que contiene una descripción equivale al significado de una oración general existencial que sólo contiene variables y términos generales -los cuales forman parte del subconjunto de las expresiones atómicas del lenguaje. En lo que respecta a éstos últimos, ofrece, a diferencia de Frege, una teoría de la referencia directa, semejante a la ofrecida por Mill a propósito de los nombres propios. Sin embargo, esta teoría adolesce del mismo defecto que la de Mill: a diferencia de la teoría descripcional propuesta para los nombres, esta teoría no involucra ninguna explicación del mecanismo por el cual un término general refiere al aspecto del mundo que refiere y no a otro. Por consiguiente, si hay alguna teoría acabada del significado en Russell, ésta no es otra que la mencionada teoría descripcional de la referencia de los nombres propios (en sentido corriente), que no hace sino darle una forma más precisa a las tesis formuladas por Frege: según Russell, el mecanismo por el cual un nombre refiere a un objeto es su asociación con una cierta descripción definida, cuya relación con el mundo debe a su vez entenderse en términos de la aplicación, por lo menos y a lo sumo a un objeto cualquiera, de ciertos términos generales; el criterio propuesto para identificar el referente de un nombre es, por tanto, la aplicación única de ciertos términos generales. En definitiva, es nuestro conocimiento de las propiedades de los individuos lo que nos permite nombrarlos -de ahí que sea posible afirmar que la concepción descripcional confiere un rol central a la dimensión epistémica. En adelante, llamaré entonces "teoría clásica" a la conjunción de las teorías descripcionales de la referencia ofrecidas sucesivamente por Frege y Russell.

1.3. Los problemas de la teoría clásica

En primer lugar, la teoría clásica no da cuenta del hecho de que distintos hablantes (o un mismo hablante en distintos momentos) asocian distintas descripciones con un término. ¿Cuál es el criterio para seleccionar a una sola de ellas como constitutiva del sentido determinante de la referencia y, por tanto, como constitutiva del significado? Volviendo al ejemplo anterior, ¿cuál es la descripción que nos da el sentido de "Tolstoi"? ¿"El autor de *Ana Karenina*", "el autor de *La guerra y la paz*", "el escritor ruso más famoso de la primera mitad del siglo XIX", "el fundador de la novela rusa moderna"? Todas estas descripciones -y muchas otras- son igualmente asociadas con el nombre en cuestión por los hablantes competentes del español.

En segundo lugar, dado que de hecho cada término es asociado con múltiples descripciones, la teoría clásica se enfrenta con un problema de ambigüedad. Cabe señalar que el propio Frege acepta esta consecuencia de su teoría, y señala que la ambigüedad es un defecto característico de los lenguajes naturales: en un lenguaje lógicamente perfecto, a cada término ha de corresponderle un único sentido -y un único referente. En otras palabras, se presenta el problema converso al tradicional problema de la identidad antes mencionado: dados los múltiples sentidos de "Charles Dodgson", se hace necesario explicar por qué

Charles Dodgson es Charles Dodgson

es una oración trivial -debería ser tan informativa como

Charles Dodgson es Lewis Carroll.

Finalmente, la teoría clásica adscribe trivialidad donde no la hay -se trata, por tanto, del problema opuesto al planteado en la objeción anterior, según la cual la teoría implica informatividad donde no la hay. Explicaré este punto en términos de nuestro ejemplo anterior: supongamos -como hemos hecho hasta el momento- que el sentido

del nombre "Tolstoi" está dado por (el sentido de) la descripción "el autor de *Ana Karenina*"; luego, la oración

Tolstoi es el autor de Ana Karenina

debería ser una oración trivial para todo aquél que comprenda el nombre; sin embargo, esto no es así: la oración en cuestión suele aportar información acerca de Tolstoi. En otras palabras, si la teoría descripcional clásica fuera cierta, dicha oración debería ser analítica (verdadera en virtud de su significado), necesaria (no podría haber sido falsa) y *a priori* (verdadera independientemente de toda experiencia), comparable por tanto a una oración como

Los solteros son no casados.

Sin embargo, esto no es así: la oración en cuestión parece ser sintética (verdadera en virtud de un hecho), contingente (podría haber sido falsa) y *a posteriori* (empíricamente verdadera). (Cabe aclarar que esta objeción está relacionada con una de las críticas de Kripke a la concepción descripcional, a saber, aquélla basada en la noción de designador rígido, que será detallada más adelante.)

Las objeciones anteriores dieron origen a una nueva versión de la teoría descripcional, distinta de la propuesta por Frege y Russell.

1.4. Las teorías "cúmulo"

Searle y Strawson han propuesto una versión refinada de la teoría clásica, la cual evidencia una clara influencia de la concepción del significado defendida por Wittgenstein en su segundo período. La concepción en cuestión presupone una nueva teoría acerca de la definición, según la cual los términos deben ser definidos no en función de un conjunto de condiciones necesarias y suficientes para su aplicación sino en función de un conjunto de características que se solapan y se entrecruzan, "una red de parecidos de familia". El ejemplo característico de Wittgenstein involucra la palabra "juego". No es posible explicitar

un conjunto de propiedades cuya sola presencia determine que algo sea un juego (como, por ejemplo, ser una actividad que involucra competencia, diversión y ejercicio) puesto que, dada una lista cualquiera de propiedades supuestamente definitorias (presumiblemente, una más completa que la anterior), alguna de las propiedades mencionadas puede faltar y sin embargo tratarse de un juego (el ajedrez no requiere ejercicio) o, por el contrario, puede ocurrir que todas ellas estén presentes y, sin embargo, que no se trate de un juego (escalar montañas es competitivo, divertido y requiere ejercicio). La palabra "juego" puede definirse en cambio en términos de una lista abierta: ninguna propiedad particular es necesaria ni ningún conjunto particular de ellas es suficiente para que a algo se le aplique la palabra "juego". El único requisito es que un subconjunto cualquiera de la lista de propiedades esté presente en cada caso. Las propiedades en cuestión, como rasgos "que se solapan y se entrecruzan", otorgan a todos los juegos un "parecido de familia".[34]

De manera semejante, según la nueva versión de la teoría descripcional, conocida como "teoría cúmulo" (*cluster theory*), el sentido de un término está constituido por el sentido no de una (única) descripción asociada sino de un conjunto o cúmulo de descripciones asociadas, obtenido a partir de una lista abierta e indefinida. De acuerdo con esto (dado que el sentido es lo que determina la referencia), un nombre refiere a aquel objeto que posee la mayor parte de las propiedades connotadas por un cúmulo de descripciones -en otros términos, a aquel objeto denotado por la mayor parte de las descripciones que in-

34. Véase Wittgenstein 1953. Considérese, por ejemplo, el siguiente comentario de Kenny, que incluye una cita de Wittgenstein: "La noción de parecido de familia se enuncia en la *Grammatik* y se desarrolla en el *Blue Book*. Las entidades que nosotros subsumimos bajo un término general, escribió Wittgenstein, no tienen porqué tener algo en común, forman una familia cuyos miembros tienen parecidos de familia. Algunos de ellos tienen la misma nariz; otros, las mismas cejas, y unos terceros, la misma manera de hablar; y estos parecidos se entrecruzan. La idea de un concepto general que sería una propiedad común de sus ejemplificaciones concretas nos conecta con otras ideas primitivas, demasiado simples, de la estructura del lenguaje" (BB 17; cf. también BB 87, 124; PI I, 67; PG 75). Kenny 1973: p. 138.

tegran el cúmulo en cuestión. Esto no quita que entre los miembros del cúmulo pueda haber diferencias: algunas descripciones son más importantes o tienen más peso que otras. Retomando nuestro ejemplo principal, en el cúmulo de descripciones asociadas con "Tolstoi", seguramente tiene más peso "el autor de *Ana Karenina*" que "el hermano del campesino Andrei". De ahí que se afirme, más estrictamente, que el sentido está constituido por (la mayor parte de) un subconjunto ponderado del cúmulo de descripciones asociadas -cierto subconjunto de descripciones muy laterales puede tener un peso 0 y, por tanto, no ser apto para dar el sentido del nombre.

Según Searle, quien sostiene esta teoría respecto de los nombres propios, todo aquél que usa un nombre debe ser capaz de contestar a la pregunta acerca de qué o quién se está hablando. La respuesta a esta pregunta involucra la especificación -y por tanto, el conocimiento- de ciertas propiedades que distinguen a ese individuo de los otros. Este conocimiento es justamente lo que se transmite al enseñar un nombre. Esto no implica que los nombres propios puedan ser reemplazados por la descripción total de las propiedades del individuo denotado, puesto que se trata de expresiones esencialmente imprecisas. En palabras de Searle:

> *Así pues la laxitud de los criterios para los nombres propios es una condición necesaria para aislar la función referencial de la descriptiva del lenguaje.*[35]

En la concepción de Searle, la diferencia entre las funciones mencionadas parece ser solamente una cuestión de grado: mientras que el nombre propio describe de una manera vaga e imprecisa, la descripción lo hace de manera inequívoca. A diferencia de Russell, Searle considera que la descripción es entonces la función semántica fundamental. La siguiente cita ilustra claramente el punto:

35. Searle 1958: p. 172.

Podemos ahora resolver nuestra paradoja: ¿tiene sentido un nombre propio? Si la pregunta se refiere a si los nombres propios son usados para describir o especificar características de los objetos, la respuesta es 'no'. Pero si la pregunta plantea si los nombres propios están lógicamente conectados con características del objeto al cual refieren, la respuesta es 'sí, de una manera laxa'.[36]

¿Cuáles son fundamentalmente las ventajas que presenta esta teoría por sobre la teoría descripcional clásica? Primeramente, la teoría cúmulo, en la medida en que reconoce que no es una única descripción sino un conjunto de ellas lo que determina el sentido de un término, logra evitar el problema planteado por la necesidad de seleccionar una única descripción entre muchas. En segundo lugar, la teoría no presenta el problema de ambigüedad antes mencionado, pues las distintas descripciones asociadas con un término forman parte de un mismo cúmulo, y es en términos de éste último y no de las primeras que se define el sentido del término. Finalmente, la teoría refinada elude el problema de la trivialidad, puesto que niega explícitamente que ninguna descripción particular sea necesariamente constitutiva del sentido de un término determinado; luego, aun cuando "el autor de *Ana Karenina*" pertenezca al cúmulo de descripciones asociadas con el nombre "Tolstoi", bien puede ocurrir que la descripción en cuestión no forme parte del subconjunto ponderado de descripciones semánticamente relevantes, por lo que la oración

Tolstoi es el autor de Ana Karenina

puede no resultar trivial, ni analítica, necesaria o a priori.

Cabe destacar, sin embargo, que las soluciones brindadas a estos problemas no son contundentes. Como podrá apreciarse a continuación, los problemas, u otros análogos a ellos, resurgen con un nuevo rostro. En primer lugar, es posible identificar un problema muy similar al mencionado en primer término: la teoría cúmulo se ve obligada a seleccionar no ya a una única pero si a algunas descripciones

36. Searle 1958: p. 173.

como integrantes del cúmulo en función del cual se define el sentido de un término. Ahora bien, no se ve claramente cuál puede ser el criterio que guía esta selección; en otras palabras, es difícil establecer un criterio que permita distinguir descripciones constitutivas de descripciones no constitutivas del significado. Dada esta dificultad, la teoría amenaza con desembocar en el holismo semántico, es decir, la tesis de que todas las descripciones asociadas, aun aquéllas que lo son de manera casual, constituyen el significado de un término. Como se verá en el capítulo siguiente, el holismo semántico es una postura implausible: por el momento, baste señalar que no es plausible pensar que todas las descripciones asociadas con un determinado término constituyen su significado, puesto que si así fuera cualquier cambio en nuestras creencias acerca del portador del término determinaría un cambio en su significado.[37]

En segundo lugar, es fácil notar que el problema de la ambigüedad persiste: cada término tendrá tantos significados cuantos cúmulos de descripciones le sean asociados. En otras palabras, no es plausible creer que todos los hablantes (ni un mismo hablante a lo largo del tiempo) asocian siempre el mismo cúmulo de descripciones con un término determinado, dado que la gente varía en sus creencias acerca de los objetos y en su consideración de cuáles de ellas son centrales y cuáles no lo son.

Finalmente, no es difícil prever el resurgimiento, bajo una forma extrema, del mencionado problema de la trivialidad. Toda oración que predique de un individuo determinado no ya una propiedad particular sino la mayor parte de las propiedades seleccionadas por el cúmulo de descripciones constitutivo del sentido del nombre del individuo en cuestión resultará trivial, desde el punto de vista de la teoría -aunque no lo será si se la deja de lado. De acuerdo con esto, una oración como

37. Más aún, dado que, desde la perspectiva holística, el significado de un cierto término depende del de todo otro en el lenguaje, las consecuencias son más drásticas: cualquier cambio, por mínimo que fuere, en el sistema de creencias determinaría un cambio en el significado del término en cuestión.

> Tolstoi tiene la mayor parte de las siguientes propiedades: autor de Ana Karenina, famoso escritor ruso del siglo XIX, uno de los fundadores de la novela moderna...

deberá ser analítica, necesaria y *a priori*, cuando claramente no parece serlo. Esta objeción será desarrollada en el siguiente apartado, en relación con la crítica de Kripke.

Por último, quisiera mencionar el problema que considero más grave, a saber, la incompletitud de la explicación ofrecida por las teorías descripcionales, en sus dos versiones. En el marco de la teoría clásica, el significado de un nombre propio es explicado en términos del significado de una descripción definida (Frege), y el significado de ésta última es a su vez explicado en términos del significado de términos generales (Russell). Pero, ¿cómo se explica el significado de los términos generales, los cuales, según Russell, refieren directamente a objetos del mundo? ¿Cuál es el mecanismo semántico en virtud del cual los términos generales refieren directamente al mundo? En este punto, no es posible encontrar explicación alguna -más allá de la mera afirmación, compartida con Mill, de que se trata de una relación directa.[38] Cabe destacar que el hecho de que se agregue que la relación en cuestión se funda en una relación de conocimiento directo no contribuye demasiado a aclarar el panorama: ¿por qué razón el conocimiento directo habría de fundamentar una relación semántica directa? A modo de ejemplo, yo puedo tener un conocimiento directo del matiz de verde que estoy viendo en este momento

38. El mismo punto puede hacerse en relación con la sola teoría de Frege en los términos siguientes. Si se toma en cuenta la interpretación antes sugerida, según la cual el sentido de un nombre puede ser identificado con un concepto descriptivo asociado con él, cabe objetar que la teoría no provee explicación alguna de cómo este concepto adquiere a su vez sentido y, en general, significado: simplemente se presupone que lo tiene. En términos de uno de los ejemplos anteriores, se dice que el significado de "Tolstoi" está constituido (en parte) por el concepto descriptivo EL AUTOR DE *ANA KARENINA*, asociado con el nombre por los hablantes competentes; sin embargo, no se explica qué es lo constituye el significado del concepto en cuestión.

y, sin embargo, expresar ese conocimiento por medio de una descripción, tal como "el matiz de verde que veo ahora". (Del mismo modo, podría pensarse que un conocimiento "por descripción" podría fundamentar una relación semántica directa.) Asimismo, en el marco de las versiones refinadas, el significado de un nombre es explicado en términos de un cúmulo de descripciones, cuyo significado es a su vez explicado en términos del significado de ciertos términos generales, el cual es a su vez explicado en términos del significado de otros términos generales: la explicación ofrecida siempre apela a las propiedades semánticas de alguna categoría lingüística.

El rasgo mencionado constituye sin duda un grave defecto de la concepción descripcional, puesto que en algún punto es necesario dar una explicación de la relación de referencia que no haga uso de nociones intencionales o lingüísticas -como la de la descripción asociada con un término. De lo contrario, la explicación ofrecida resulta circular: descansa en una supuesta comprensión (de ciertas aplicaciones) del concepto que se intenta explicar -a saber, la referencia.

Así como los problemas semánticos tradicionales ponían en graves aprietos a la teoría milleana de la referencia directa, los problemas mencionados en este apartado socavan seriamente a la concepción descripcional. Es el momento entonces de considerar a un nuevo enfoque, que históricamente surge como una alternativa más sólida frente a los problemas planteados por las teorías descripcionales.

2. Kripke o el origen del modelo causal

2.1. Designadores rígidos y causalidad

Kripke es uno de los primeros filósofos que presentan una concepción alternativa de cómo un término refiere a un objeto, la cual puede caracterizarse en términos generales como "concepción causal": de acuerdo con ésta, un término refiere a un objeto en virtud

no de una descripción asociada sino de hallarse relacionado causalmente de la manera apropiada con aquél.[39] Ahora bien, dado que la propuesta de la nueva concepción está estrechamente relacionada con la crítica de la concepción descripcional, la exposición consta de dos partes principales. La primera contiene la crítica al modelo descripcional y, por tanto, la fundamentación de la tesis negativa de que los nombres y los términos generales no son semánticamente equivalentes a las descripciones. En la segunda parte, me dedico a reconstruir el núcleo central de la tesis positiva ofrecida, a saber, la mencionada tesis según la cual dichos términos refieren a los objetos del mundo en virtud de mantener apropiadas relaciones causales con ellos. Vale la pena destacar que las tesis de Kripke comprenden tanto a los nombres propios como a los términos generales de clases naturales -en adelante, me referiré a ellos como "términos generales".[40]

2.1.1. La propuesta negativa: nombres y términos generales vs. descripciones

Así como la teoría descripcional refinada (cluster theory) une inextricablemente el significado de los nombres y los términos generales al significado de las descripciones, Kripke sostiene que no hay razones para creer que tal vínculo sea real; por el contrario, la intuición misma muestra que uno y otro están muy alejados. En su opinión, el mecanismo por el cual tanto un nombre propio como un término general refieren a un objeto es muy distinto del mecanismo por el cual una descripción lo hace; los nombres y los términos generales funcionan semánticamente de una manera muy distinta de

39. Quienes presentaron casi simultáneamente teorías del mismo tipo son fundamentalmente Donnellan y Putnam. Véanse Donnellan 1966, 1972, 1989. La teoría de Putnam será examinada en el capítulo siguiente.
40. Cabe aclarar que la teoría también se considera válida para ciertos términos que designan fenómenos naturales, tales como "luz", "calor", "sonido" y, muy probablemente, para los adjetivos correspondientes, tales como "luminoso", "caliente", "ruidoso".

como lo hacen las descripciones. Los primeros, a diferencia de éstas últimas, refieren a sus objetos de manera directa, esto es, sin involucrar la adscripción de propiedad alguna al objeto referido. Por consiguiente, no tiene sentido alguno valerse, como hacen las teorías descripcionales, de las descripciones para explicar el significado de los nombres y los términos generales. La teoría de la referencia para éstos ha de ser entonces, según Kripke, una teoría de la referencia directa. Ahora bien, ¿qué argumentos ofrece Kripke en favor de esta tesis central? En mi opinión, los argumentos ofrecidos pueden ser clasificados en dos grupos: por un lado, están aquéllos que dependen esencialmente de la llamada "tesis de los designadores rígidos"; por otro, hay un argumento que es independiente de aquélla.

2.1.1.1. La tesis de los designadores rígidos y el argumento modal-epistemológico

Comenzaré por introducir la definición kripkeana de "designador rígido": un designador es rígido si y sólo si designa o refiere al mismo objeto en todos los mundos posibles en los que ese objeto existe.[41] El análisis de esta definición nos involucrará momentáneamente con el uso de ciertos conceptos metafísicos, tales como el de necesidad y el de esencia, que no son el objeto central de este trabajo; por consiguiente, cabe aclarar que sólo me ocuparé de ellos

41. La definición dada en el texto explicita la noción general de designador rígido dada por Kripke, que es la que se presupondrá a lo largo de la discusión. Pero puede ser conveniente aclarar que, como distingue Salmon, hay dos tipos de designadores rígidos: (i) los *persistentes*, esto es, aquellas expresiones que designan al mismo objeto en todo mundo posible en el que ese objeto existe y que no designan nada en aquellos mundos posibles en los que ese objeto no existe y (ii) los *obstinados*, es decir, aquellas expresiones que designan al mismo objeto en todos los mundos posibles, independientemente de que el objeto en cuestión exista o no. Los designadores rígidos que son, a la vez, obstinados y persistentes son denominados "designadores *fuertemente rígidos*", esto es, expresiones que designan rígidamente a un objeto que existe en todos los mundos posibles. Según Salmon, el sentido de la noción más frecuentemente aludido por Kripke es el primero, a saber, el de designador rígido persistente. Véase Salmon 1982: cap. 1, pp. 32-41.

en la medida en que sea pertinente hacerlo para el examen de las tesis semánticas. No es mi propósito entonces examinar en detalle la relación entre la semántica y la metafísica kripkeanas sino presentar su concepción semántica, para lo cual me involucraré con la metafísica sólo en la medida de lo necesario.

La idea general de Kripke es la siguiente: los nombres propios y los términos generales son designadores rígidos; las descripciones definidas, en cambio, no son, por lo general, designadores rígidos (con excepción de las descripciones matemáticas, como "la raíz cuadrada de 9", que constituyen un subconjunto muy pequeño dentro del conjunto de descripciones expresables en el lenguaje natural); luego, el significado de los nombres y los términos generales no puede ser explicado en términos del significado de las descripciones.

La afirmación de que los nombres y los términos generales, a diferencia de las descripciones, son designadores rígidos equivale a la afirmación de que unos y otras tienen distintas propiedades modales: los nombres y términos generales permiten identificar a los objetos a pesar del cambio de sus propiedades, es decir, en la totalidad del espectro de los mundos posibles en los que esos objetos existen; las descripciones, en cambio, sólo permiten identificarlos en función de sus propiedades cambiantes, por tanto, sólo en aquellos mundos posibles en donde los objetos conservan las propiedades en cuestión. A modo de ejemplo, "Tolstoi" permite referir a Tolstoi en todos los mundos posibles en donde Tolstoi existe, aun cuando se trate de mundos en donde aquél no escribe *Ana Karenina* o carece de algunas de sus otras propiedades; "el autor de *Ana Karenina*", en cambio, sólo permite referir a Tolstoi en aquellos mundos posibles en donde aquél escribe *Ana Karenina*.

Ahora bien, una primera objeción surge espontáneamente ante esta propuesta: ¿es acaso plausible creer que uno está hablando de Tolstoi cuando se refiere a un individuo que no escribe *Ana Karenina*? En general, ¿puede un individuo y, en general, un objeto seguir siendo el mismo aun cuando le falten muchas de sus propiedades importantes? En otras palabras, si se dejan de lado tales propiedades, ¿qué es lo que determina la identidad de un objeto a través del espectro de los mundos posibles? La respuesta de Kripke a esta

pregunta comienza, en mi opinión, con su énfasis en el hecho de que la concepción propuesta, lejos de requerir extraños criterios de identidad a través de los mundos posibles, tiene una fuerte base intuitiva.

A fin de iluminar la intuición subyacente, es preciso tener en cuenta la concepción kripkeana de los mundos posibles -concepto estrechamente relacionado con el de designador rígido. Kripke no concibe a los mundos posibles a la manera de Lewis, es decir, como mundos reales pero no efectivos en donde se encuentran las contrapartes de los objetos efectivamente existentes.[42] Para Kripke, los mundos posibles son situaciones contrafácticas o historias posibles del mundo, constituidas por los mismos objetos que constituyen el mundo efectivo. En opinión de Kripke, su concepción está más cerca de la intuición que la de Lewis, puesto que, cuando uno describe situaciones contrafácticas, de hecho se refiere -o es intuitivo pensar que así ocurre- a los objetos efectivamente existentes y no a sus fantasmales contrapartes. A modo de ejemplo, cuando uno se pregunta qué habría pasado si Napoleón no hubiera invadido Rusia en 1812, está especulando sobre la historia posible del Napoleón y la Rusia efectivos y no de sus contrapartes. De acuerdo con esto, la tesis de que los nombres y términos generales son designadores rígidos debe entenderse como la tesis, fuertemente intuitiva, de que los nombres y términos generales son instrumentos que nos permiten hablar de los objetos reales efectivamente existentes en situaciones contrafácticas (es decir, situaciones en las que carecen de sus propiedades habituales) y especular sobre lo que podría haberles ocurrido.

En segundo lugar, Kripke reconoce que la tesis semántica de los designadores rígidos está estrechamente relacionada con cierta concepción del esencialismo metafísico, a saber, la tesis que afirma la existencia de esencias o propiedades esenciales. De acuerdo con esta concepción, hay ciertas propiedades, denominadas "esenciales", que pertenecen necesariamente a los objetos (es decir, que les pertenecen en todos los mundos posibles en donde aquéllos existen); ahora bien,

42. Véase Lewis 1986.

son las propiedades esenciales las que constituyen el correlato ontológico de los designadores rígidos. Como puede preverse, no es fácil especificar cuáles son estas propiedades. Según Kripke, en el caso de las personas, la esencia está constituida por el óvulo y el espermatozoide que le dieron origen en cada caso; en el caso de los objetos materiales, apela no sólo al origen sino también a la materia de la que están compuestos; en el caso de los tipos naturales, la esencia está constituida por la microestructura física.

De acuerdo con esto, un individuo podría carecer de la mayor parte de las propiedades usualmente consideradas importantes -como la de ser el autor de *Ana Karenina* en el caso de Tolstoi o ser el responsable de la invasión de Rusia en 1812 en el caso de Napoleón- y seguir siendo el mismo, pero no sería el mismo si su origen fuera distinto. Con la sola excepción de la propiedad de tener un cierto origen, todas las propiedades objetivas, inclusive las consideradas importantes, resultan ser entonces accidentales. De este modo, afirmar que los nombres son designadores rígidos equivale a afirmar que los nombres refieren a esencias individuales, las cuales están constituidas exclusivamente por aquellas propiedades que determinan el origen (o la materia) de los individuos. En términos de nuestro ejemplo, afirmar que "Tolstoi" es un designador rígido equivale a afirmar que "Tolstoi" designa al individuo que posee la propiedad esencial de ser el hijo del señor y la señora Tolstoi, el cual no podría haber sido otro que Tolstoi mismo, mientras que afirmar que "el autor de *Ana Karenina*" no es un designador rígido equivale a afirmar que "el autor de *Ana Karenina*" designa al individuo que posee la propiedad accidental de ser el autor de *Ana Karenina*, el cual podría ciertamente no haber sido Tolstoi.

Del mismo modo, un tipo natural podría carecer de la mayor parte de las propiedades usualmente consideradas importantes, es decir, las propiedades fenomenológicas -como la de ser amarillo en el caso del oro o ser transparente en el caso del agua- y seguir siendo el mismo, pero no sería el mismo si su microestructura física fuera distinta. Con la sola excepción de la propiedad de tener una cierta microestructura física, todas las propiedades objetivas, inclusive las propiedades fenomenológicas, resultan ser entonces accidentales. De

este modo, afirmar que los términos generales son designadores rígidos equivale a afirmar que refieren a esencias naturales, las cuales están constituidas exclusivamente por aquellas propiedades que determinan la microestructura física de los objetos. En términos de uno de los ejemplos anteriores, afirmar que "agua" es un designador rígido equivale a afirmar que "agua" designa al tipo natural que posee la propiedad esencial de tener la composición química H_20, el cual no podría haber sido otro que el agua misma, mientras que afirmar que "la sustancia transparente que llena los ríos y los lagos" no es un designador rígido equivale a afirmar que "la sustancia transparente que llena los ríos y los lagos" designa al tipo natural que posee las propiedades accidentales de ser transparente y llenar los ríos y los lagos, el cual podría ciertamente no haber sido el agua.

En lo que sigue, haré una breve reflexión sobre la relación entre la tesis de los designadores rígidos y el esencialismo metafísico. En general, se suele interpretar que la relación propuesta por Kripke entre las dos instancias mencionadas es la de implicación lógica: más específicamente, se piensa que Kripke considera que la tesis de los designadores rígidos implica lógicamente el esencialismo metafísico. Ante todo, quiero consignar que no estoy completamente segura de que esta interpretación sea correcta: pienso que efectivamente Kripke considera que la relación entre las tesis en cuestión es muy profunda o, más estrictamente, conceptual; pero tengo algunas dudas acerca de si realmente le otorga a la tesis semántica prioridad por sobre la tesis metafísica. Dejando de lado el punto de qué sea lo que efectivamente sostiene Kripke al respecto, me interesa criticar entonces lo que sin duda se ha interpretado que dice: en mi opinión, no es posible considerar que la tesis de los designadores rígidos implica el esencialismo metafísico. Mis razones para sostener esto son las siguientes. En primer lugar, considero que no hay fundamentos suficientes para creer que nuestra concepción del lenguaje, esto es, una parte muy pequeña del mundo, puede fundamentar nuestra concepción del mundo en general; en otras palabras, en mi opinión, no es posible pensar en ningún caso que las generales y abarcativas tesis metafísicas se derivan de tesis locales como las semánticas; *a fortiori*, el esencialismo no puede derivarse de la tesis de los

designadores rígidos. Sin pretender profundizar en este punto, lo que sin duda excedería los límites de este trabajo, quiero tan sólo señalar que las cuestiones metafísicas, en virtud de su carácter general, deberían ser consideradas prioritarias respecto de las cuestiones semánticas.[43] De este modo, lo que considero que debe afirmarse es que la tesis de los designadores rígidos *presupone* el esencialismo. Ahora bien, una vez explicitado el orden de la relación (primero está el esencialismo y después, la tesis de los designadores rígidos), es preciso aclarar que, en mi opinión, no se trata de una relación lógica o conceptual sino de una relación *empírica*, en otras palabras, considero que el argumento que permite pasar de la tesis esencialista a la tesis de los designadores rígidos no es un argumento deductivo sino un argumento abductivo o inferencia a la mejor explicación: dada la adopción del esencialismo, la mejor explicación del funcionamiento semántico de los nombres propios y los términos generales involucra la defensa de la tesis de los designadores rígidos. Finalmente, cabe destacar que el esencialismo presupuesto por Kripke no es cualquier tipo de esencialismo sino una versión particular del mismo, a saber, el compromiso con la existencia de esencias constituidas ya sea por las propiedades que determinan el origen de un objeto individual, ya sea por las propiedades que determinan la microestructura física de una clase natural. Este último punto puede apreciarse claramente si se tiene en cuenta que hay una versión distinta del esencialismo, a saber, aquélla según la cual las esencias individuales están constituidas por la mayor parte de las propiedades comúnmente atribuidas a los individuos mientras que las esencias de las clases naturales lo están por la mayor parte de las propiedades a ellas atribuidas, que resulta plenamente compatible no con el modelo kripkeano sino con la concepción descripcional. Tres puntos merecen entonces ser destacados respecto de la relación entre las tesis semántica y metafísica aquí consideradas: (i) no es correcto fundamentar el esencialismo metafísico en la tesis de los designadores rígidos sino que, a la inversa,

43. He intentado defender la tesis propuesta acerca de la relación entre la semántica y la metafísica en mi tesis doctoral. Argumentos similares pueden encontrarse en Devitt 1991b y Barrio 1998.

ésta última depende de la primera; (ii) la relación entre ambas tesis no es la relación lógica de implicación sino la relación de inferencia a la mejor explicación o, en otras palabras, el argumento que permite pasar de una tesis a otra no es un argumento deductivo sino abductivo; como mencioné anteriormente, ese argumento podría ser reconstruido en los siguientes términos: dado el esencialismo, la mejor explicación del funcionamiento semántico de los nombres es aquélla que apela a la noción de designador rígido; y (iii) no es el esencialismo metafísico en general sino cierta versión particular de la tesis esencialista lo que constituye un presupuesto kripkeano; del mismo modo, podría considerarse que la concepción descripcional descansa sobre una versión distinta del esencialismo metafísico.

El mismo Kripke parece suscribir esta última sugerencia, como puede considerarse que se pone de manifiesto en uno de sus argumentos específicos en contra de la concepción descripcional. Para presentar este argumento, conocido como "argumento modal", utilizaré un nuevo ejemplo:

(i) Si la teoría descripcional es verdadera, entonces Mussorgsky posee necesariamente la mayor parte de las propiedades mediante las que usualmente se lo caracteriza -ser el autor de *Khovanshchina*, ser el más famoso compositor ruso del siglo XIX, ser el mejor compositor del Grupo de los Cinco, etc. -en otras palabras, las propiedades en cuestión constituyen su esencia.

(ii) Mussorgsky no posee necesariamente ningún conjunto tal de propiedades, puesto que es posible concebir un mundo en donde Mussorgsky no posee ninguna de ellas y sin embargo sigue siendo Mussorgsky -en otras palabras, las propiedades en cuestión no constituyen su esencia.

Luego (iii) la teoría descripcional no es verdadera.

Téngase en cuenta que de la versión refinada de la teoría descripcional, según la cual un nombre propio refiere al individuo denotado por la mayor parte de las descripciones de un cúmulo con él asociado, se sigue que el individuo en cuestión ha de poseer la disyunción inclusiva de la lista total de propiedades que

la comunidad de hablantes cree que posee o comúnmente le atribuye -no podría haber ninguna situación contrafáctica en la que no poseyese ninguna de esas propiedades. De acuerdo con esto, es posible ofrecer una versión más perspicua del argumento anterior en los siguientes términos:

(i) Si la teoría descripcional es verdadera, entonces "Mussorgsky es o bien el autor de *Khovanshchina* ...o bien el mejor compositor del Grupo de los Cinco", donde los puntos suspensivos reemplazan a la enumeración de todas las propiedades que comúnmente se le atribuyen, es una oración necesariamente verdadera.

(ii) "Mussorgsky es el autor de *Khovanshchina* ...o bien el mejor compositor del Grupo de los Cinco", donde los puntos suspensivos reemplazan a la enumeración de todas las propiedades que comúnmente se le atribuyen, no es una oración necesariamente verdadera.

Luego, (iii) la teoría descripcional no es verdadera.

Dada la unión tradicional entre los conceptos de necesidad, aprioridad y analiticidad, es posible reconstruir argumentos paralelos al anterior en términos de estas dos últimas nociones. Ejemplificaré solamente uno de ellos, conocido también como "argumento epistemológico". (Cabe aclarar que, en realidad, como sugiere el título de este apartado, es posible considerar que se trata de un solo argumento que puede denominarse "modal-epistemológico".)

(i) Si la teoría descripcional es verdadera, entonces "Mussorgsky es o bien al autor de *Khovanshchina* ... o bien el mejor compositor del Grupo de los Cinco", donde los puntos suspensivos reemplazan a la enumeración de todas las propiedades que comúnmente se le atribuyen, es una oración verdadera *a priori*.

(ii) "Mussorgsky es el autor de *Khovanshchina*... o bien el mejor compositor del Grupo de los Cinco", donde los puntos suspensivos reemplazan a la enumeración de todas las propiedades que comúnmente se le atribuyen, no es una oración verdadera *a priori*.

Luego, (iii) la teoría descripcional no es verdadera.

Sobre la base de lo anterior, es claro que las segundas premisas de estos argumentos se basan en la noción de designador rígido: la oración mencionada en cada una de ellas no es ni necesariamente verdadera ni verdadera *a priori* porque el nombre "Mussorgsky" es un designador rígido. Veamos de qué manera una cosa se sigue de otra: si "Mussorgsky" es un designador rígido, entonces, en otro mundo posible, designará a la misma persona, Mussorgsky, que designa en este mundo; sin embargo, en ese mundo, esa persona tal vez no tenga ninguna de las propiedades mencionadas en los cúmulos asociados: podría no haber compuesto *Khovanshchina*, podría no integrar el Grupo de los Cinco, etc; como resultado de ello, la oración mencionada en (ii), lejos de ser trivial, expresa un hecho contingente y empírico acerca de Mussorgsky, discernible a partir del estudio histórico y revisable a la luz de futuros descubrimientos. En este punto, es posible apreciar claramente de qué manera Kripke otorga nuevo fundamento a la objeción basada en la atribución de trivialidad que se mencionó en el apartado anterior.

Como puede preverse, estos mismos argumentos pueden ser reconstruidos en relación con el uso de términos generales.[44] En términos de la estructura anterior,

(i) Si la teoría descripcional es verdadera, entonces "Un tigre es o bien un animal cuadrúpedo... o bien un animal amarillo tostado con rayas negras transversales y panza blanca", donde los puntos suspensivos reemplazan a la enumeración de todas las propiedades que comúnmente se le atribuyen, es una oración necesariamente verdadera y verdadera *a priori*.

(ii) "Un tigre es o bien un animal cuadrúpedo ... o bien un animal amarillo tostado con rayas negras transversales y panza blanca", donde los puntos suspensivos reemplazan a la enumeración de todas las propiedades que comúnmente se le atribuyen, no es ni una oración necesariamente verdadera -puesto que bien podría haber tigres que no tuvieran ninguna de esas propiedades y seguirían, sin

44. Véase Kripke 1980: p. 26.

embargo, siendo tigres- ni una oración verdadera *a priori* -puesto que, para establecer su verdad, es preciso observar cómo son de hecho los tigres.

Luego, (iii) la teoría descripcional no es verdadera.

Por consiguiente, la fuerza de los argumentos anteriores depende en gran parte de la concepción kripkeana de los nombres y los términos generales como designadores rígidos y las intuiciones modales que subyacen a esta concepción. Cabe destacar, sin embargo, que la aceptación del argumento no implica necesariamente un compromiso con una ontología de mundos posibles: como sugiere el propio Kripke, es posible interpretar el discurso acerca de los mundos posibles como una manera metafórica de hablar de posibilidades y necesidades o, en otros términos, como modos de conceptualizar el mundo efectivo. De todos modos, especialmente si se desconfía de la legitimidad de la apelación a intuiciones modales, es preciso tener en cuenta que Kripke propone otro argumento, conocido como "argumento semántico", que es totalmente independiente de aquéllas.

2.1.1.2. El argumento semántico

Se trata, en mi opinión, del argumento más importante ofrecido por Kripke en contra de la concepción descripcional. Por medio de él, Kripke ataca al núcleo de las teorías descripcionales refinadas, es decir, el hecho de que, en la medida en que el significado de todo nombre propio y todo término general está determinado por (el significado de) un conjunto de descripciones asociadas, el hablante competente en el manejo de aquéllos ha de tener una creencia verdadera o conocimiento acerca del objeto referido; más específicamente, la tesis de que comprender un nombre propio o un término general equivale a conocer hechos que permiten identificar su referente. Según Kripke, este conocimiento, atribuido al hablante por la concepción descripcional, no es ni necesario ni suficiente. El argumento puede reconstruirse en los siguientes términos:

(i) Si la teoría descripcional es verdadera, entonces el hablante competente en el manejo de un nombre propio/término general no

puede ser ignorante ni estar equivocado acerca del objeto referido por el término en cuestión.

(ii) En muchas ocasiones, el hablante es ignorante o sólo tiene creencias falsas acerca del objeto referido -lo que no obstaculiza en absoluto la referencia.

Luego, (iii) la teoría descripcional no es verdadera.

La premisa (i) expresa, como vimos, lo que puede ser considerado el núcleo central de toda teoría descripcional. La premisa (ii), que constituye la clave del argumento, es fundamentada por Kripke mediante la presentación de distintos tipos de evidencia. A continuación, clasificaré esa evidencia en diferentes grupos.

En primer lugar, el autor considera que (a) hay muchas ocasiones en las que el hablante competente en el manejo de un nombre no cree que las descripciones asociadas seleccionen un único individuo. El ejemplo que da Kripke involucra el nombre "Cicerón": la mayor parte de los hablantes competentes sólo atribuyen al individuo denotado por este nombre la propiedad de ser un orador romano, de la cual no creen que pertenezca sólo a Cicerón.

En segundo lugar, (b) hay ocasiones en las que no sólo el hablante no cree que la descripción asociada seleccione un único individuo sino que ésta de hecho no lo hace, puesto que o bien selecciona a varios, como es el caso de "un orador romano" antes mencionado, o bien no selecciona a ninguno, como es el caso de todas las descripciones asociadas con el nombre "Jonás", tales como "el profeta que fue tragado por un enorme pez", "el profeta que fue tragado por una ballena", "el profeta que fue a Nínive a predicar", etc, ninguna de las cuales selecciona al individuo referido por el nombre.

En tercer lugar, Kripke destaca que (c) hay ocasiones en las que el hablante no asocia descripción alguna con un determinado nombre, como tal vez sea el caso de "Einstein", si se dejan de lado descripciones muy generales tales como "un científico famoso" -en cuyo caso se trataría de un nuevo ejemplo de (a).

De acuerdo con (a), (b) y (c) es posible tener un alto grado de ignorancia o estar completamente equivocado acerca del individuo referido, y, sin embargo, usar el nombre para referirse a ese individuo.

El conocimiento de propiedades identificadoras no es por tanto una condición necesaria para la referencia de un nombre.

Por último, Kripke señala que (d) hay ocasiones en las que el hablante asocia cierta descripción con un nombre, la cual selecciona efectivamente un único individuo, pero éste no es el referente del nombre en cuestión. Por ejemplo, supongamos que no fue Beethoven quien compuso la 7ª. sinfonía sino un músico desconocido, Fritz Weiss, quien se la regaló a Beethoven y le pidió que guardara el secreto y dijera al mundo que era obra suya; bajo este supuesto, "Beethoven" refiere a Beethoven y "el autor de la 7ª. sinfonía", asociada por el mundo entero con el nombre anterior, refiere a Fritz Weiss -y por más que insistamos en asociarlos no lograremos que refieran al mismo individuo, en particular, no lograremos que el nombre "Beethoven" refiera al individuo seleccionado por la descripción, esto es, Fritz Weiss.[45]

De acuerdo con (d), la asociación de una descripción identificadora ("el autor de la 7ª. sinfonía") con un nombre ("Beethoven") no puede ser tampoco una condición suficiente para que el nombre en cuestión refiera a un individuo determinado (Fritz Weiss) y no a otro (Beethoven).

Cabe destacar que los mismos tipos de evidencia pueden ser recolectados para probar una tesis paralela respecto de los términos generales: a saber, es posible tener un alto grado de ignorancia o estar completamente equivocado acerca de las propiedades de la clase natural referida, y, sin embargo, usar el término general para referirse a la clase en cuestión. El conocimiento de propiedades identificadoras tampoco es por tanto una condición necesaria para la referencia de un término general. Esto puede ser apreciado en fragmentos como el siguiente:

> *(...) usamos 'oro' como un término para una cierta clase de cosa. (...) La clase de cosa es pensada como si tuviera ciertas marcas que*

45. Como puede apreciarse, este ejemplo es similar al presentado por Kripke en Kripke 1980, pp. 91-2, que involucra a los nombres "Gödel", "Smith" y a la descripción "el autor del teorema de incompletitud".

la identifican; algunas de estas marcas pueden no ser realmente verdaderas del oro. Podríamos descubrir que nos equivocamos con respecto a ellas.[46]

De este modo, el argumento semántico es también conocido como "argumento basado en la ignorancia y el error". Lo que me interesa destacar es que este argumento objeta la denominada "dimensión epistémica del significado", introducida por la teoría fregueana del sentido y preservada por las teorías descripcionales posteriores. Desde el punto de vista de Kripke, los nombres propios y los términos de clases naturales no tienen tal dimensión: a diferencia de las descripciones, aquéllos refieren a los objetos directamente, sin adscribirles ninguna propiedad, es decir, sin describirlos. La relación entre tales tipos de términos y sus correspondientes objetos no está mediatizada por ningún concepto descriptivo. En este punto, cabe entonces destacar que la crítica de Kripke a la concepción descripcional involucra una reivindicación parcial de la vieja teoría de Mill: por un lado, tanto para Kripke como para Mill, los nombres propios son expresiones puramente denotativas; por otro, Kripke, a diferencia no sólo de la concepción descripcional sino también de Mill, considera que los términos de clases naturales también lo son.[47]

En síntesis, la crítica de Kripke a la concepción descripcional puede entenderse sobre la base de la distinción entre los conceptos de dar el significado y fijar la referencia de un término. Según Kripke, el error de la concepción descripcional consiste básicamente en haber asignado ambas funciones a las descripciones -esto tiene un claro origen en la noción fregueana del sentido de un nombre, dimensión esencial de su significado, definida en términos de(l sentido de) una descripción asociada y cuya función semántica fundamental es, como puede recordarse, la de determinar el referente. En su opinión, en cambio, las descripciones asociadas pueden servir para fijar la referencia de los términos pero esto no implica que formen parte de su

46. Kripke 1980: p. 125. Los destacados son míos.
47. Véase, por ejemplo, Kripke 1980: pp. 133-4.

significado. En este punto, es útil mencionar el famoso ejemplo de la definición del metro patrón: un metro es la longitud de la barra de París.[48] La descripción "la longitud de la barra de París", asociada en virtud de esta definición con la expresión "un metro", bien puede ser utilizada en un momento dado para fijar el referente de "un metro", como ocurre en la definición anterior; pero esto no implica que "la longitud de la barra de París" y "un metro" tengan el mismo significado, puesto que la primera refiere a la medida, cualquiera sea ésta, que satisface esa descripción (mañana, si la barra es calentada, tal vez se trate de una medida más grande que la de hoy; es concebible un mundo en donde esa longitud sea distinta) mientras que la segunda refiere a una medida determinada (que no varía ni en este mundo a través del tiempo ni a través de los mundos posibles). En otros términos, la descripción refiere a una medida en virtud de connotar cierta propiedad -ser la longitud de la barra de París- que la medida en cuestión debe poseer; el nombre refiere a una medida directamente. Ahora bien, es el momento de introducir la tesis positiva de Kripke, esto es, la tesis que explica en qué consiste esa relación de referencia directa. Es en función de esta propuesta positiva que la teoría de Kripke se distingue radicalmente de la de Mill.

2.1.2. La propuesta positiva: un esbozo del modelo causal

Según Kripke, el mecanismo por el cual un nombre propio y un término de clase natural refieren a un objeto está determinado por la existencia de una relación causal entre ambos. Las relaciones causales, en la medida en que constituyen relaciones físicas, parecen ofrecer un camino de salida del círculo intencional, y otorgarle así al significado una explicación plenamente compatible con el fisicalismo, es decir, la concepción ontológica general según la cual todos los

48. Véase Kripke 1980: pp. 62-3. Cabe aclarar, para disipar desconciertos, que Kripke es consciente de haber elegido un ejemplo extraño, en el sentido de que la expresión "un metro" no constituye un típico ejemplo de los denominados "nombres propios".

hechos -químicos, biológicos, psicológicos, semánticos, etc- son explicables en términos naturales o, en otras palabras, son reductibles, en sentido amplio, a hechos físicos.[49] De este modo, el enfoque causal puede ser interpretado como un claro intento por naturalizar la referencia. Ahora bien, Kripke distingue un mecanismo de fijación de la referencia y un mecanismo de transmisión de la misma, en los cuales, como se verá a continuación, la causalidad interviene de distintas maneras.

El mecanismo de fijación opera en el momento hipotético en que el nombre es usado por primera vez para hacer referencia a un individuo. Para caracterizar ese momento, Kripke utiliza la expresión "bautismo inicial": a partir de una relación perceptiva directa con un objeto determinado, una cierta expresión queda causalmente "anclada" o "fundada" en el objeto en cuestión. Es este vínculo causal fundacional lo que da al término su significado, o, en otras palabras, lo que determina el objeto para cuya referencia va a usarse a partir de entonces. Cabe aclarar, sin embargo, que en muchos casos el bautismo inicial es realizado con la ayuda de descripciones, las cuales contribuyen a la identificación del objeto referido. Pero, como se aclaró anteriormente, esto no debe conducir a la idea de que las descripciones en cuestión constituyen el significado del término. Las descripciones son instrumentos útiles a la hora de fijar la referencia, especialmente, cuando la ostensión no es posible; sin embargo, no forman parte del significado del término -por las razones antes consignadas.[50] A modo de ejemplo, (un caso de) el nombre "Lucio" está fundado en mi hijo porque en un momento dado, mi marido y yo hemos expresado, señalando al bebé, "Llamémoslo 'Lucio'". Alternativamente, hubiéramos podido exclamar: "El bebé precioso que están viendo se llama 'Lucio'" o "Vamos a llamar 'Lucio' al bebé que va a nacer", pero no por ello "Lucio" habría de ser semánticamente equivalente a las descripciones "el bebé precioso que están viendo" ni "el bebé que va a nacer".

49. El sentido amplio de "reducción" al que me refiero es aquél según el cual la superveniencia de todos los hechos en los hechos físicos es una forma de reducción. Para el concepto de superveniencia, véase, por ejemplo, Kim 1982.
50. Véase, por ejemplo, Kripke 1980: p. 103, nota 42.

Del mismo modo, un término general designa rígidamente una clase natural si y sólo si designa rígidamente una cierta muestra de una clase natural que mantiene la relación de ser de la misma clase con otras muestras en todos los mundos posibles.[51] De este modo, un término general, si bien se funda originariamente en una muestra particular, está causalmente relacionado con todas las muestras de la misma clase natural. A modo de ejemplo, el término "agua" está inicialmente fundado en una muestra particular de agua -la cual mantiene con todas las otras muestras la relación de ser de la misma clase. Ahora bien, muy probablemente, en el bautismo hipotético de una clase natural, se utilice alguna descripción, tal como "la sustancia instanciada por las cosas que están ahí o al menos por casi todas ellas". Sin embargo, la descripción en cuestión, si bien contribuye indudablemente a fijar el referente del término, no forma parte de su significado. En términos más próximos a los de Kripke, las propiedades accidentales y, por tanto, contingentes que se mencionan característicamente en descripciones como la anterior sirven para fijar la referencia de un término pero no para dar el significado.

El denominado "mecanismo de transmisión de la referencia" es considerado, en cambio, puramente causal: tanto los nombres propios como los términos de clases naturales son transmitidos a los miembros de la comunidad lingüística que no estuvieron presentes en los bautismos por medio de una cadena causal de comunicación. En otras palabras, es en virtud del contacto causal con los miembros de la comunidad lingüística partícipes del bautismo cómo aquéllos que no lo fueron adquieren la habilidad de usar y comprender los usos ajenos de un nombre o un término general. Existe entonces una cadena causal muy compleja que relaciona cada caso de éstos con los objetos con los que fueron inicialmente vinculados. Para retomar ejemplos anteriores, los padres transmiten la referencia de "Lucio" a otras personas cuando comentan a un amigo por teléfono

51. Como la relación de designación involucrada es la relación de designación rígida -es decir, es una relación válida en todo mundo posible-, la relación de ser de la misma clase comprende a muestras de todos los mundos posibles -por lo que constituye una relación trasmundana.

"Te dejo porque Lucio tiene hambre y voy a darle de comer" o cuando le dicen a algún vecino poco amistoso "Lamentamos mucho que Lucio haya llorado tanto anoche: le están saliendo los dientes". Del mismo modo, se supone que la referencia de "agua" o de "oro" son transmitidas, de grupo en grupo y de generación en generación, mediante una cadena causal. En términos de Kripke:

> *Aún más importante, el nombre de la especie puede pasarse de eslabón en eslabón, exactamente como en el caso de los nombres propios, de manera que quienes han visto muy poco o ningún oro pueden sin embargo usar el término. Su referencia se determina mediante una cadena causal (histórica), no mediante el uso de ningún ejemplar.*[52]

Me interesa destacar entonces un punto que ha sido señalado anteriormente, a propósito del argumento semántico: se trata de una nueva concepción acerca de cómo un signo se relaciona con un objeto que no asigna rol alguno al conocimiento que el hablante pueda tener del objeto en cuestión. En los términos anteriores, el significado carece de dimensión epistémica: no hay sentido mediador sino sólo referencia o referencia directa. Por consiguiente, el sentido deja de ser, como lo era en el marco de las teorías descripcionales, el punto de contacto entre la teoría del significado y la teoría de la comprensión. Desde el nuevo punto de vista, en la medida en que el significado no involucra sentido alguno, la comprensión no implica conocimiento: como se puso de manifiesto en la exposición del mecanismo de transmisión, comprender un signo equivale a insertarse adecuadamente en la cadena causal que conduce al referente. El concepto de cadena causal de comunicación es también un concepto clave en la teoría de la comprensión asociada al nuevo modelo causal.

De acuerdo con esto, Kripke rechaza no sólo la tesis tradicional de que el uso de nombres propios y términos generales requiere el conocimiento de propiedades identificadoras del objeto referido sino también una nueva sugerencia de Strawson según

52. Kripke 1980: p. 145.

la cual la transmisión de la referencia requiere el conocimiento de aquél de quien se toma prestado el término.[53] De acuerdo con esto, la comprensión de "Mussorgsky", por ejemplo, exige su asociación con una descripción tal como "el individuo a quien Stravinsky llama 'Mussorgsky'"; a su vez, la comprensión de "Stravinsky" exige su asociación con una descripción tal como "el individuo a quien Andrei llama 'Stravinsky'" y así sucesivamente hasta llegar a un nombre de cuyo referente se posea conocimiento identificador. Según Kripke, si bien es distinta de la explicación típica de la concepción descripcional, esta propuesta también exige conocimiento donde puede no haberlo: es posible que el hablante haya olvidado por completo o esté completamente equivocado acerca de la fuente del término y sea, sin embargo, capaz de usarlo correctamente.

Finalmente, cabe destacar que el rol central desempeñado por las cadenas causales de comunicación da a esta teoría semántica un carácter histórico: se trata de una teoría que se ocupa de las relaciones causales que de hecho tienen lugar entre nombres e individuos, por un lado, y entre términos generales y clases naturales, por otro, a lo largo de la historia de la comunidad lingüística.[54] Es preciso aclarar, sin embargo, que la propuesta de Kripke no constituye una teoría acabada en todos sus detalles sino tan sólo un esbozo de teoría, cuyos detalles requieren mayor elaboración -como se pondrá claramente de manifiesto si se examinan detenidamente los problemas que presenta.[55]

2.2. Los problemas de una teoría causal de la referencia directa

2.2.1. El problema del respecto

Ante todo, la teoría causal está afectada por un problema de ambigüedad denominado "problema del respecto" (*qua problem*). De

53. Véase Kripke 1980: pp. 96-8.
54. Véase, por ejemplo, Kripke 1980: pp. 101-2.
55. Véase, al respecto, Kripke 1980: p. 101.

acuerdo con éste, la relación causal no parece ser suficiente para fijar la referencia: admite demasiados candidatos como posibles referentes de un término; en otras palabras, en la medida en que relaciona a cada término con múltiples aspectos de un objeto, la relación causal tiene múltiples extremos ontológicos. En consecuencia, si el significado del término está constituido exclusivamente por el objeto referido, el término en cuestión resulta ambiguo. Este problema puede apreciarse mejor por medio de un ejemplo: la existencia de un vínculo causal fundacional entre "Balzac" y una cierta librería, basado en la percepción directa del lugar, nos permite considerar a "Balzac" como el nombre de la librería en cuestión; pero ¿sobre qué base es posible seleccionar esa relación en lugar de la relación, basada en el mismo acto perceptivo, entre "Balzac" y (por ejemplo) la puerta del negocio o el cartel grande que está en la vidriera? En otras palabras, ¿cómo hacer para seleccionar un poder causal y dejar de lado a los demás?

En lo que respecta a los términos de clases naturales, como se destacó anteriormente, si bien se fundan originariamente en una muestra particular, se los considera relacionados con todas las muestras de la misma clase natural. Este último punto pone de relieve el problema que nos ocupa: cada fragmento de materia es una muestra de muchas clases de cosas, por lo que mantiene la relación de igualdad de clase con muchas otras muestras. ¿Es acaso posible seleccionar, en términos de la noción de causalidad, un aspecto determinado de la muestra como el único pertinente para el significado de un cierto signo? En otras palabras, ¿es acaso posible determinar, a la luz de la relación causal, qué naturaleza de la muestra es semánticamente pertinente? Supongamos que "x" ha sido fundado en una muestra de agua: la muestra en cuestión es también una muestra (entre otras cosas) de la clase de las cosas transparentes y de la clase (no natural) de los líquidos que pueden beberse en el almuerzo. ¿Cómo puede la mera existencia de una relación causal fijar el significado de "x"? La respuesta es simple: no puede, puesto que, sobre la base de la sola relación causal, es preciso concluir que el signo "x" refiere tanto al agua como a la clase de las cosas transparentes y a la clase de los líquidos que pueden beberse en el almuerzo.

Parecería que la única manera de seleccionar una relación causal entre muchas, y por consiguiente, la única manera de fijar un referente para nombres y términos generales, es tener en cuenta las intenciones de los miembros de la comunidad lingüística presentes en el bautismo. Es porque un grupo de personas es capaz de describir de alguna manera a la causa de su experiencia perceptiva que "Balzac" refiere a un negocio (y no a una puerta ni a un cartel); en otras palabras, es porque las personas en cuestión intentan referirse al objeto entero (y no a sus partes) que "Balzac" refiere a un negocio (y no a una puerta ni a un cartel). Del mismo modo, es porque las personas poseen ciertas categorías conceptuales, como por ejemplo la de clase natural, y desean referirse al agua que cierto grafisma, "x" o en español "agua", refiere al agua (y no, por ejemplo, a la clase de los líquidos que pueden beberse en el almuerzo). En ambos casos, es claro que hay ciertas habilidades cognitivas en juego. Si esto es así, la noción de causalidad no parece ser suficiente para determinar la referencia y explicar, de esa manera, el significado: a fin de evitar el problema del respecto, requiere ser complementada con nociones intencionales, como las de intención y deseo. Pero éstas últimas son actitudes proposicionales, las cuales, como se verá más adelante, pueden ser concebidas como distintos tipos de relaciones con oraciones, compuestas a su vez por términos. De este modo, se está explicando la referencia de un término en función de otros términos; en otras palabras, en la explicación de la referencia de los términos se da por sentado que los términos tienen referencia. El riesgo de ser circulares y violar la condición fisicalista parece entonces socavar seriamente a las teorías histórico-causales.

En este punto, vale la pena hacer algunas aclaraciones. En primer lugar, es importante destacar que el mencionado problema de ambigüedad es un problema semántico, referido a la naturaleza del significado. Para reproducirlo en pocas palabras: si se pretende definir el significado de los términos exclusivamente en términos de relaciones histórico-causales con objetos, debe afrontarse el hecho de que aquéllos resultan ambiguos, puesto que las relaciones causales en cuestión tienen múltiples correlatos ontológicos. Ahora bien, este problema puramente semántico no debe ser confundido con el problema, en

apariencia similar, señalado por Quine como fundamento de su tesis de la inescrutabilidad de la referencia.[56] Según Quine y, en la misma línea, Davidson, para comprender un lenguaje extraño -y, paralelamente, para construir una teoría acerca de la comprensión de un lenguaje extraño-, es necesario basarse en ciertas consideraciones epistémicas, en particular, la evidencia del asentimiento -o la disposición a asentir- a oraciones en circunstancias públicamente observables; pero la evidencia en cuestión, aun en el caso ideal, no determina unívocamente la referencia de las expresiones suboracionales; por consiguiente, una misma asignación de condiciones veritativas a una oración dada es compatible con la adscripción de distintos esquemas referenciales. En otros términos, la evidencia del asentimiento a oraciones, si bien permite al traductor de un lenguaje establecer las condiciones veritativas de las mismas, no le permite determinar unívocamente a qué refieren los términos singulares del lenguaje en estudio ni de qué son verdaderos sus predicados. La ostensión misma no parece tener ninguna utilidad al respecto:

> *Señala un conejo y has señalado el estadio temporal de un conejo, una parte (fundamental) del conejo, la fusión (de todas las partes) del conejo, y el fragmento espacio-temporal donde la conejidad se manifiesta.*[57]

Como es sabido, esta tesis, denominada "tesis de la inescrutabilidad de la referencia", es una de las premisas fundamentales a partir de las cuales Quine construye el argumento de la indeterminación de la traducción, según el cual no hay nada en la realidad que permita considerar a una de las traducciones ofrecidas como la traducción correcta, es decir, no hay nada objetivo que pueda considerarse que las distintas traducciones ofrecidas pretenden reflejar con mayor o menor éxito.

56. Véanse Quine 1960: cap. 2 y Davidson 1977.
57. Quine 1960: cap. 2, pp. 52-53.

Si bien el problema de Quine es en algún sentido similar al problema de la ambigüedad -dado que es, en definitiva, un problema de ambigüedad-, hay un aspecto en el que uno y otro difieren claramente. El problema del que parte Quine es un problema básicamente gnoseológico o epistemológico: a la luz de la evidencia conductística, es decir, el único tipo de evidencia de la que puede disponer, el traductor de un lenguaje extraño es incapaz de establecer a qué refieren sus términos o, dicho de otro modo, la observación de la conducta lingüística no provee al traductor de evidencia suficiente como para poder atribuir relaciones referenciales objetivas a los términos de un cierto lenguaje; como consecuencia de ello, toda traducción está esencialmente indeterminada. Dado su escepticismo general acerca de la existencia de significados y su concepción general del lenguaje como un conjunto de prácticas o disposiciones a la conducta verbal (concepción que, como se mencionó en la introducción, es heredada o compartida por Davidson), Quine no duda en concluir a partir del mencionado problema gnoseológico acerca de las limitaciones del traductor, una tesis semántica acerca de la naturaleza del significado mismo: la referencia de los términos de un lenguaje es inescrutable -esto es, la mencionada tesis de la inescrutabilidad de la referencia. Pero no hay que olvidar que el problema originario es un problema gnoseológico, relacionado con la tarea de traducción y el tipo de evidencia disponible al traductor, que es fundamentalmente ajeno al modelo causal.

Nótese, además, que la tesis semántica concluida por Quine es una tesis general y profunda: la referencia está absolutamente indeterminada, no hay modo de seleccionar una relación referencial entre las múltiples posibles. El rechazo de la causalidad se desprende de esta tesis: el término "causalidad" no tiene un referente unívoco o, en otras palabras, no hay criterio objetivo alguno que permita seleccionar una interpretación entre las múltiples posibles de "causalidad". Los múltiples correlatos ontológicos de un término son los correlatos de múltiples relaciones -entre las que cabría incluir a la causalidad tal como se la entiende comúnmente. El problema semántico que se presenta a los teóricos causales no es, por el contrario, un problema tan general o tan profundo: la ambigüedad no es lo que impide que se

acepte a la causalidad como determinante de la referencia sino que se presenta una vez que se la ha aceptado, como consecuencia de ello. Los múltiples correlatos ontológicos de un término son todos ellos correlatos de la relación causal. El problema aquí no es la imposibilidad de apelar a la noción de causalidad para definir la referencia sino el hecho de que el concepto en cuestión no parece ser suficiente.

En segundo lugar, tal vez podría pensarse que la tesis de los designadores rígidos y la peculiar concepción esencialista a ella unida permiten a la teoría causal de Kripke evitar el problema del respecto. Según esta idea, el correlato ontológico de un nombre o de un término general está unívocamente determinado por cuanto está constituido por una determinada esencia (ya sea individual o natural). En mi opinión, sin embargo, esto no es así: independientemente de la ontología defendida o presupuesta -es decir, independientemente del tipo de entidades con las cuales queramos comprometernos-, siempre va a existir una multiplicidad de instancias superpuestas en el extremo ontológico de la relación causal. El problema del respecto es el problema de cómo justificar cuál de ellas es semánticamente pertinente. En términos de nuestros ejemplos anteriores, es necesario justificar porqué "Mussorgsky" refiere a la esencia de una persona y no a la esencia de una cara o de una nariz y por qué "agua" refiere a H_2O y no a la clase no natural de los líquidos que pueden beberse en el almuerzo.

Ahora bien, hechas estas aclaraciones, cabe preguntarse si Kripke ha tenido en cuenta este problema en alguna medida. El único lugar en donde se hace referencia al problema discutido es en la nota 58 del libro de Kripke, a propósito de una discusión con Geach acerca de las propiedades esenciales. Kripke sintetiza una crítica de Geach en estos términos:

> *De acuerdo con Geach, dado que todo acto de señalar es ambiguo, quien bautiza un objeto señalándolo tiene que aplicar una propiedad genérica* (sortal property) *para desambiguar su referencia y garantizar criterios correctos de identidad a través del tiempo; por ejemplo, alguien que asigne una referencia a "Nixon" señalándolo tiene que decir "Uso 'Nixon' como un nombre de ese hombre", haciendo*

así a un lado las tentaciones por parte de quien lo escucha de considerar que está señalando a una nariz o a una rebanada de tiempo. La propiedad genérica es entonces, en algún sentido, parte del significado del nombre: los nombres tienen después de todo un sentido (parcial), aun cuando sus sentidos puedan no ser suficientemente completos para determinar sus referencias, como tampoco lo son en la teoría descripcionista y en la teoría del cúmulo de descripciones.

El problema aludido por Geach es claramente nuestro problema del respecto. Kripke es entonces consciente de él, pero no parece otorgarle la importancia debida, como se pone de manifiesto en su respuesta a la objeción, parte de la cual encuentro ilustrativo transcribir a continuación:

(...) hay claramente un hueco sustancial entre premisa y conclusión. Pocos hablantes aprenden de hecho la referencia de un nombre dado mediante una ostensión, e incluso si adquieren el nombre mediante una cadena de comunicación que lleva en último término a una ostensión, ¿por qué habría de ser de algún modo parte del "sentido" del nombre la propiedad genérica supuestamente usada en la ostensión? No se da ningún argumento aquí.[58]

En mi opinión, esto representa un intento infructuoso por invertir la carga de la prueba; en otras palabras, es Kripke y no Geach quien debe probar algo, a saber, por qué las descripciones utilizadas en la fijación de la referencia de un término no constituyen su "sentido" y, por tanto, su significado. Mis razones para creer esto son las siguientes. Kripke admite explícitamente que para fijar la referencia es necesario hacer uso de descripciones; es decir, las descripciones tienen para él un rol clave en el mecanismo de fijación de la referencia. Ahora bien, ¿por qué negar entonces que en la instancia del bautismo inicial aquéllas forman parte del significado del término involucrado? En otras palabras, ¿por qué no reconocer que (i) hay dos mecanismos distintos por los cuales un término refiere a un objeto: uno de ellos,

58. Kripke 1980: pp. 122-3, nota 58.

característico del bautismo inicial, es descriptivo-causal y el otro, propio de los usos posteriores, es puramente causal y (ii) ambos mecanismos referenciales son constitutivos del significado, por lo que, en definitiva, las descripciones utilizadas en el bautismo inicial forman parte del significado del término? Creo que esto es lo que debería concluirse si se parte de la base, común a todas las teorías referencialistas del significado, de que el significado de un nombre está constituido, por lo menos en parte, por aquello que determina su referencia. Dado este presupuesto general, el intento kripkeano por separar la teoría del significado de la teoría de la fijación de la referencia me parece *ad hoc*. En otras palabras, hacer esto último parece presuponer el abandono del presupuesto general en cuestión o, más específicamente, el abandono del propósito de que la explicación propuesta acerca de la referencia logre iluminar la naturaleza del significado. Pero, entonces, ¿cuál podría ser el propósito de ocuparse de la referencia? Desde mi punto de vista, el interés por la referencia sólo puede basarse en el interés por explicar el significado.

No es implausible creer que lo que conduce a Kripke a separar de una manera tajante sentido o significado, por un lado, y referencia, por otro, hasta el punto de afirmar que lo que contribuye a fijar la referencia no es en modo alguno constitutivo del significado, es su suscripción de la tesis de los designadores rígidos. La razón de esta conjetura reside en el hecho de que la tesis en cuestión afirma precisamente esa separación tajante, sobre la base de intuiciones modales e independientemente de la consideración de los mecanismos explicativos del uso lingüístico.

2.2.2. El problema del sentido

Dejando de lado el problema de la ambigüedad, cabe destacar que esta teoría, en tanto concibe a los nombres y a los términos generales como expresiones puramente denotativas, reaviva los fantasmas de los problemas semánticos tradicionales. En otros términos, si el significado de tales términos es identificado con el objeto denotado o referido, ¿cómo se evitan problemas tales como el problema de la identidad, la paradoja de la denotación y el problema de las oraciones acerca de

entidades ficticias, entre otros? Dados éstos últimos, las teorías de la referencia directa, como la de Kripke, no parecen tener muchas probabilidades de éxito. Cabe pensar que Kripke, enfrentado con estos problemas, opta por lo que puede denominarse "una estrategia de exportación".[59] De acuerdo con esta tesis, los problemas en cuestión son exportados o trasladados fuera del ámbito de la semántica, sobre la base de que no constituyen problemas genuinamente semánticos. Kripke en particular considera que no están relacionados con la naturaleza del significado sino con los procesos de adscripción de actitudes proposicionales, esto es, creencias, deseos, intenciones, etc; por consiguiente, en su opinión, no afectan a la semántica sino a cualquiera sea la disciplina que se ocupa de la adscripción de actitudes proposicionales -tal vez la epistemología, la psicología o la filosofía de la mente.

Ahora bien, desde mi punto de vista, esto constituye un error: la estrategia de exportación es altamente cuestionable. Si bien es cierto que, debido al escaso grado de desarrollo de la semántica científica, el objeto de estudio de la semántica no está claramente delimitado (como lo están, por ejemplo, los de la física, la lingüística y la economía), hay ciertas intuiciones acerca de lo que compete y no compete a la semántica. En particular, es intuitivo pensar que los procesos de adscripción de actitudes proposicionales le competen en la medida en que pueden tener influencia en la determinación de la naturaleza del significado; más específicamente, es intuitivo pensar que ésta última está estrechamente relacionada con los propósitos perseguidos mediante los procesos de adscripción de actitudes. Veamos con algún detalle la relación entre ambas instancias.

Una oración adscriptora de una actitud proposicional, tal como

Thales cree que todo procede del agua

consiste en una oración del lenguaje natural mediante la cual se adscribe una cierta actitud proposicional, en este caso, una creencia,

59. Véase Kripke 1979. Véase, por ejemplo, Devitt 1996: pp. 179-186.

a un individuo determinado; la creencia en cuestión es identificada o individualizada mediante el uso de otra oración (oración adscripta), que es semánticamente equivalente a la oración subordinada que aparece en la oración adscriptora, en este caso,

Todo procede del agua.

Las actitudes proposicionales se individualizan en función de dos parámetros: por un lado, el modo de la actitud involucrada, es decir, si se trata de una creencia, un deseo, una intención, etc; por otro, el significado de la oración adscripta. En este punto, cabe hacer entonces dos aclaraciones. En primer lugar, este criterio de individualización de las actitudes proposicionales presupone una concepción relacional de las mismas, esto es, una concepción según la cual tener una determinada actitud proposicional es estar relacionado de cierta manera con una cierta oración. En segundo lugar, es posible considerar que la oración con la cual está relacionado el sujeto de una actitud es o bien una oración del lenguaje natural o bien una oración del lenguaje del pensamiento: en cualquiera de los dos casos, se considera que se trata de una oración semánticamente equivalente a la oración subordinada incluida en la oración adscriptora. En este trabajo, presupondré la concepción relacional, pero no tomaré partido respecto del segundo punto, es decir, dejaré abierta la cuestión acerca de si la oración adscripta, correspondiente a la oración subordinada de la oración adscriptora, pertenece al lenguaje natural o al lenguaje mental. Todo lo que se dice es, en principio, neutral respecto de esta cuestión.[60]

Ahora bien, el solo hecho de que la distinción entre actitudes proposicionales se haga en parte en función de lo adscripto en cada caso y lo adscripto sea siempre una oración, esto es, una entidad con significado, puede conducir a pensar que los procesos de adscripción

60. Para un análisis de distintas concepciones acerca de las actitudes proposicionales y una defensa de la concepción relacional y el punto de vista de que las oraciones con las que se relaciona el sujeto de actitudes proposicionales son oraciones del lenguaje del pensamiento, véase, por ejemplo, Fodor 1978.

de actitudes competen claramente a la semántica. Sin embargo, la tesis anterior depende de la creencia en la existencia de un vínculo aun más estrecho: los propósitos perseguidos al realizar adscripciones (por ende, al emitir oraciones como "Thales cree que todo procede del agua") influyen en la naturaleza de lo adscripto (esto es, en la naturaleza del significado de oraciones del lenguaje natural, como "Todo procede del agua" u oraciones del lenguaje mental, como TODO PROCEDE DEL AGUA).[61] Más específicamente, si mi propósito al realizar la adscripción antes mencionada es transmitir lo que cree Heráclito acerca del mundo, entonces será útil hacer uso de una oración cuyo significado involucre algún tipo de relación con el mundo -en otros términos, de una oración cuyo significado sea explicado en términos representacionales. De acuerdo con esto, los significados que necesitamos son aquéllos que responden a nuestros propósitos semánticos.

Cabe aclarar que estas consideraciones involucran cierto cambio en la manera de concebir "las cuestiones relativas a la adscripción". Por lo general, cuando se tratan tales cuestiones, como ocurre en el tratamiento de los problemas semánticos tradicionales por parte de Frege y Russell, el problema planteado es el de establecer el significado de las oraciones adscriptoras, tales como la antes mencionada "Thales cree que todo procede del agua". Mi intención, en cambio, es ocuparme de tales oraciones sólo en la medida en que puede considerarse que arrojan luz sobre los significados de las oraciones adscriptas -tales como la igualmente mencionada "Todo procede del agua"-; en otras palabras, como destaqué anteriormente, mi intención es partir de las adscripciones para desentrañar la naturaleza de los significados adscriptos. De ahí que no me interese por establecer,

61. En adelante, omitiré la distinción entre unas y otras y me referiré a las oraciones en general mediante el procedimiento de entrecomillarlas. No debe pensarse por ello que he tomado partido por la concepción relacional de las actitudes proposicionales que concibe al extremo oracional de la relación en términos de oraciones del lenguaje natural. Mi intención, como señalé en el texto principal, es permanecer neutral en este punto; sin embargo, omitir la distinción en cuestión evita complicar innecesariamente la exposición.

a la manera de Frege, el significado que adquieren ciertos términos cuando se encuentran en contextos opacos, sino los distintos tipos de significados (referentes, sentidos, etc) para términos que en general es preciso distinguir en función de los propósitos e intereses a los que obedecen los procesos de adscripción.

Ahora bien, la pregunta clave es: ¿cuáles son nuestros propósitos semánticos, es decir, con qué propósitos nos adscribimos actitudes proposicionales unos a otros? Y aun, de modo más amplio, ¿con qué propósitos adscribimos significados tanto a los estados mentales como a las emisiones lingüísticas de nuestros semejantes?

Uno de los propósitos que pueda considerarse que es satisfecho mediante la adscripción de significados es la expansión o el incremento de nuestro conocimiento acerca del mundo. De acuerdo con esto, la siguiente oración adscriptora de creencia

Juan cree que va a llover

unido al hecho de que confío en Juan en lo que concierne al pronóstico del tiempo, me permiten inferir que efectivamente va a llover -lo cual involucra un incremento de mi conocimiento acerca del mundo. Ahora bien, el cumplimiento del propósito en cuestión implica que el significado adscripto ha de tener ciertos rasgos: ¿cuáles son ellos?[62] Como puede preverse, la respuesta representacionalista es la siguiente: los términos del lenguaje efectivamente representan o refieren a los objetos del mundo porque es el rasgo representacional o referencial el único que permite explicar el cumplimiento o éxito de nuestro propósito de incrementar nuestro conocimiento acerca del mundo por medio de la adscripción de actitudes proposicionales y, en general, de significados. En términos de nuestro ejemplo, dado que la adscripción antes mencionada me permite adquirir cierta información acerca de cierto

62. Esto abre la posibilidad de que los rasgos en cuestión hayan sido inicialmente seleccionados en virtud de su aptitud para cumplir con el mencionado propósito, es decir, que el fundamento del rol semántico de tales rasgos esté dado por el hecho de que son esos rasgos, y no cualesquiera otros, los que permiten cumplir con los propósitos semánticos.

aspecto del mundo (determinada por la inferencia de que va a llover), entonces es plausible afirmar que "(va a) llover" refiere a lluvia (futura).

Como es sabido, se suele considerar que otro de los propósitos que se persiguen comúnmente mediante la adscripción de significados es la explicación de la conducta. De este modo, la adscripción anterior, permite asimismo explicar(se) la acción por parte de Juan de tomar prestado un paraguas y su disposición a ponerse el piloto antes de salir: nuevamente, dado que la adscripción en cuestión es útil a nuestro propósito semántico, es plausible afirmar que "(va a) llover" refiere a lluvia (futura). Sin embargo, hay muchos casos en los que la adscripción de significados identificados en función de los referentes de los términos no es suficiente para llevar a cabo la explicación deseada. A modo de ejemplo, la decisión de Pedro de no asistir a un determinado concierto puede ser explicada mediante un caso de

Pedro cree que tocará el primer violinista del cuarteto local

pero no mediante un caso de

Pedro cree que tocará el reciente ganador del premio Bach

aun cuando "el primer violinista del cuarteto local" y "el reciente ganador del premio Bach" refieran al mismo individuo: el punto es que, a los fines de la explicación de la conducta, parece contar no sólo la relación con el individuo referido (la referencia) sino también el modo en que esa relación tiene lugar (el modo de referir a un cierto individuo, esto es, lo que Frege denominó "sentido"), y en lo que a éste último respecta las expresiones anteriores pueden ser consideradas distintas entre sí. Esta vez es entonces preciso señalar lo siguiente: dado que la primera emisión, "Pedro cree que tocará el primer violinista del cuarteto local", pero no la segunda, "Pedro cree que tocará el reciente ganador del premio Bach", es útil a nuestro propósito de explicarnos la conducta de Pedro, entonces es plausible afirmar no sólo que "el primer violinista del cuarteto local" refiere a cierto individuo en particular sino también que lo hace de determinada manera o bajo determinado modo.

Como se habrá advertido, las consideraciones anteriores sugieren que los propósitos que nos guían en la adscripción de significados permiten clasificar las oraciones adscriptoras o, para decirlo más brevemente, las adscripciones de actitudes proposicionales de una manera por demás conocida, a saber, en adscripciones transparentes o *de re* y adscripciones opacas o *de dicto*.[63] Las adscripciones transparentes o *de re* involucran una relación referencial directa con un objeto específico; en otros términos, son adscripciones de referentes. Las adscripciones opacas o *de dicto*, en cambio, atribuyen una relación referencial con un objeto bajo un modo específico de presentación del objeto en cuestión; en otros términos, son adscripciones de modos de presentación o sentidos. Es preciso enfatizar, sin embargo, que mi objetivo presente no es profundizar en el análisis de la distinción anterior (para establecer, por ejemplo, si involucra o no una clasificación exhaustiva de las adscripciones) sino senalar el punto general indicado más arriba: la existencia de una relación muy estrecha entre los propósitos que guían la adscripción de significados, por un lado, y la naturaleza de los significados adscriptos, por otro. Más específicamente, es plausible afirmar que, dados (por lo menos algunos de) los propósitos que de hecho se persiguen y se cumplen mediante la adscripción de significados, los significados adscriptos deben tener ciertos rasgos; aún más específicamente, dado que mediante la adscripción de significados de hecho se intenta y se logra (por lo menos) incrementar nuestro conocimiento del mundo y explicar la conducta, entonces es plausible concluir que los significados adscriptos deben ser identificados no sólo en términos de referentes o individuos referidos sino también en términos de modos de presentación de los mismos -en otras palabras, que los significados tienen una naturaleza dual. Si esto es así, no parece correcto afirmar, a la manera de Kripke, que los problemas que afectan al proceso de adscripción de significado son ajenos a la semántica, esto es, a la ciencia que se ocupa de la naturaleza de los significados. Más aún, tampoco parece correcto afirmar, como sugiere el autor, que lo que

63. El *locus classicus* en donde se explicita claramente esta distinción es Quine 1956.

fue caracterizado como "problema del sentido" sea un problema que concierne a la adscripción de significado y no al significado mismo.

En síntesis, parecería que adscribimos aquellos significados que sirven a nuestros propósitos semánticos -o que la naturaleza de nuestros significados es exactamente aquélla que nos permite cumplir con nuestros propósitos semánticos. Si esto es así, es decir, si los significados que de hecho adscribimos son aquéllos que nos permiten cumplir con nuestros propósitos semánticos, me interesa destacar los siguientes puntos. En primer lugar, es posible adoptar una cierta metodología semántica que consiste en partir de la observación de las adscripciones corrientes de significado para luego proponer una hipótesis acerca de la naturaleza de los significados adscriptos. En segundo lugar, es posible concluir que no es correcto, como sugiere Kripke, excluir de la semántica la consideración de los procesos de adscripción y sus distintos propósitos. Finalmente, es posible afirmar que el problema del sentido es un problema genuinamente semántico, concerniente a la naturaleza del significado, cuya resolución exige la consideración de los mencionados propósitos.

Capítulo 2
Las teorías mixtas

Dentro del marco histórico-causal, es posible distinguir dos tipos de teorías mixtas que representan intentos radicalmente distintos por superar el problema del sentido, es decir, dos tipos de teorías que no ejemplifican la estrategia de exportación mencionada en el capítulo anterior. Por un lado, están las teorías del doble factor, encabezadas por la teoría de Putnam, que proponen el agregado de un nuevo factor semántico, el estereotipo o rol conceptual. Dado que el nuevo factor es conceptualmente distinto e independiente de la referencia, se trata de teorías semirrepresentacionalistas del significado, es decir, como se explicitó anteriormente, teorías que se comprometen con la existencia de propiedades semánticas que no contribuyen a determinar la referencia. Por otro lado, está la teoría descriptivo-causal, formulada por Devitt y Sterelny, que intenta dar una solución no sólo al problema del sentido sino también al problema del respecto. Esta teoría también reconoce la existencia, junto a la referencia, de otra propiedad semántica: la de referir a un objeto de determinado modo. Pero, a diferencia de los estereotipos o roles conceptuales, los modos no son independientes de la referencia; se trata, por tanto, de una teoría representacionalista del significado, es decir, una teoría que no se compromete ontológicamente con propiedades semánticas que no contribuyan a determinar la referencia. Los

modos postulados por la teoría descriptivo-causal se distinguen, sin embargo, de los sentidos fregueanos por no ser descriptivos sino puramente causales: están constituidos no por descripciones o cúmulos de descripciones asociadas sino por los tipos de cadenas causales que conectan a los términos con sus referentes. Ahora bien, mi objetivo en este capítulo ha sido exponer y criticar ambos enfoques, de modo tal de dejar en claro mi posición sobre el tema. La teoría defendida como consecuencia de ello es de tipo descriptivo-causal, pero se diferencia de la propuesta por Devitt y Sterelny en lo que respecta a la concepción del sentido: en mi opinión, los sentidos no pueden ser concebidos como modos puramente causales. Por lo demás, considero, a diferencia de estos autores, que el subconjunto de las representaciones básicas no puede ser explicado en términos histórico-causales: a fin de dar cuenta de aquél, la teoría descriptivo-causal aquí defendida debe ser complementada mediante una teoría causal de otro tipo, no histórico -del que me ocuparé en el capítulo siguiente.

1. Putnam y las teorías del doble factor

1.1. El aporte de Putnam[64]

1.1.1. En torno a la Tierra Gemela

Como es sabido, Putnam presenta un nuevo argumento, distinto de los ofrecidos por Kripke, en contra de las teorías descripcionales,

64. Cabe señalar que la teoría de Putnam expuesta en este capítulo puede ser considerada una versión (especial, en virtud del carácter semirrepresentacionalista mencionado en el texto) de la teoría correspondentista de la verdad y fue sostenida por el autor durante lo que se conoce como el primer período de su reflexión filosófica. Posteriormente, en un segundo período, Putnam se transformó en el crítico más acérrimo de la teoría en cuestión y defendió, consistentemente, una teoría distinta que, lejos de ejemplificar la concepción correspondentista, constituye un ejemplo de teoría semántica basada en la noción de uso o condiciones de asertabilidad. La presente teoría se halla expuesta paradigmáticamente en Putnam 1975; véanse Putnam 1978 para la transición y Putnam 1981 y 1988 para la nueva teoría.

basado en el conocido experimento mental de la Tierra Gemela.[65] El experimento nos incita a imaginar a dos personas molecularmente idénticas -Oscar 1 y su gemelo Oscar 2- que habitan mundos idénticos en todas sus características -la Tierra y la Tierra Gemela respectivamente- con una sola excepción: el agua de la Tierra Gemela no tiene la composición química H_2O sino XYZ. Ahora bien, dado que tanto el agua de la Tierra como la de la Tierra Gemela tienen las mismas propiedades observables o fenomenológicas, Oscar 1 y Oscar 2 asocian con el término "agua" las mismas descripciones: "líquido transparente, sin sabor y sin olor", "sustancia que llena los ríos y los lagos", "sustancia que sacia la sed", etc; en otras palabras, Oscar 1 y Oscar 2 tienen los mismos estados mentales (que incluyen a las antes mencionadas actitudes proposicionales, tales como las creencias, las intenciones y los deseos) respecto del agua. El punto de Putnam es que, a pesar de ello, no es posible considerar que sus respectivas emisiones de la palabra "agua" tengan el mismo significado, puesto que en el caso de Oscar 1 "agua" refiere a H_2O, mientras que en el caso de Oscar 2 "agua" (o su equivalente en el español gemelo) refiere a XYZ. En síntesis, no hay nada acerca de los estados internos de Oscar 1 y Oscar 2, ninguna capacidad descriptiva o identificatoria, que permita seleccionar H_2O en lugar de XYZ o viceversa; sin embargo, "agua" en boca de Oscar 1 refiere a H_2O y "agua" en boca de Oscar 2 refiere a XYZ –por lo que no es posible considerar que ambos usos tienen el mismo significado.

La conclusión que infiere Putnam es que, a diferencia de lo que sostienen las teorías descripcionales, la asociación de descripciones con un término determinado no es suficiente para determinar el referente, y por tanto, el significado del término en cuestión -nuevamente, Oscar 1 y Oscar 2 asocian las mismas descripciones con el término "agua" y, sin embargo, este término tiene para cada uno de ellos un referente, y por tanto, un significado distinto (dado que el referente es por lo menos una parte del significado). En otras palabras, hay por lo menos una dimensión del significado, la referencia,

65. Véase Putnam 1975.

que no depende de la asociación de un término con otros sino de su relación directa con las entidades del mundo que nos rodea.[66] Por consiguiente, la referencia y, por tanto, el significado no están determinados por los estados internos o psicológicos de los hablantes. En términos del famoso *slogan* de Putnam: "meanings' just ain't in the head", es decir, los significados no están en la cabeza.[67]

Cabe aclarar que si bien el argumento de la Tierra Gemela involucra explícitamente sólo a los términos generales de clases naturales (términos como "agua", "oro", "tigre"), considero que podría ser extendido sin problemas a los nombres propios, de los cuales Putnam no se ocupa en absoluto. A modo de ejemplo, cuando Oscar 2 dice

Beethoven es el mejor compositor de la historia

el nombre "Beethoven" por él utilizado no refiere a Beethoven sino al Beethoven Gemelo; del mismo modo, la historia mencionada es la historia de la Tierra Gemela; las descripciones asociadas con los términos en cuestión, y por ende, sus estados psicológicos, son, sin embargo, idénticos a los de Oscar 1 cuando éste se refiere a Beethoven y la historia de la Tierra por medio de una oración que suena exactamente como la anterior. La razón por la que Putnam se centra casi exclusivamente en los términos generales es que su intención es, en última instancia, fundamentar el realismo científico; de este modo, constituye sin duda un exponente de la extendida adopción, en la teoría semántica, de las llamadas "estrategias ontológicas": en otras palabras, es uno de los muchos filósofos que a lo largo de este siglo han intentado fundamentar una postura metafísica -en este caso, el realismo científico- en una tesis semántica -en este caso, el rechazo de la teoría descripcional de la referencia para los términos de clases naturales.[68]

66. Burge establece una tesis similar que hace hincapié en la influencia no del entorno físico sino del entorno social. Véanse Burge 1979, 1986 y 1988.
67. Putnam 1975: p. 225.
68. El argumento en favor del realismo científico que parte del rechazo de la teoría descripcional de la referencia puede reconstruirse suscintamente en los siguientes términos: si la referencia de los términos generales que com-

Sobre la base del argumento presentado, Putnam propone entonces explicar la referencia de los términos de clases naturales por medio de la combinación de los dos componentes principales que vimos a propósito de la teoría de Kripke: la tesis de los designadores rígidos y el modelo histórico-causal. Por un lado, según Putnam, el carácter de designador rígido de los términos generales se pone de manifiesto en las definiciones ostensivas de tales términos. Para seguir con el ejemplo, el término "agua" se define ostensivamente de la siguiente manera:

Esto (señalando un líquido determinado) es agua

En esta definición, el demostrativo "esto" funciona como un designador rígido: agua es cualquier cosa que esté en la relación transmundana de ser el mismo líquido con la muestra de líquido referido mediante "esto" en el mundo actual. En términos levemente más formales,

(Para todo mundo posible M) (para todo x de M) (x es agua si y sólo si x está en la relación de ser el mismo líquido con el objeto al que "esto" -en la oración anterior- refiere en el mundo actual)

---

ponen las leyes científicas no depende del conjunto de descripciones asociadas en el marco de una cierta teoría, entonces puede considerarse que no varía o es preservada a través del cambio de teorías; de este modo, es posible pensar que las teorías científicas propuestas a lo largo de la historia son teorías *distintas* acerca de los *mismos* aspectos de un mundo que existe independientemente de aquéllas. Se presupone entonces que la existencia de una relación de correspondencia entre el lenguaje de las teorías científicas y un mundo de entidades inobservables permite fundamentar la existencia objetiva e independiente de lo mental de éste último, esto es, la tesis central del realismo científico –o realismo acerca de las entidades inobservables postuladas por la ciencia. Como mencioné anteriormente a propósito del análisis de la teoría de Kripke, considero que la adopción de estrategias ontológicas en semántica no es adecuada, por las razones antes esbozadas. Una defensa de este punto puede encontrarse, por ejemplo, en Devitt 1991b, en Barrio 1998 y en mi tesis doctoral.

Afirmar que los términos generales son designadores rígidos equivale a afirmar que dichos términos tienen un componente indicador oculto, esto es, que, aunque no parezca, su funcionamiento semántico es semejante al de las expresiones indicadoras, tales como "esto". Estas últimas constituyen, por excelencia, un tipo de expresiones para el cual es obvio que no vale el principio general de que la referencia está determinada por un conjunto de descripciones asociadas: es claro que el objeto externo forma parte *directamente* de su significado.

En cuanto al recurso al modelo causal, Putnam propone la hipótesis de la división del trabajo lingüístico. Esta última es semejante a la división del trabajo entre los distintos miembros de una comunidad: así como no todos los oficios y profesiones son accesibles a cada uno de los miembros de una comunidad, tampoco los criterios para la aplicación de todos los términos generales son accesibles a todos los miembros de la comunidad lingüística. Ciertas distinciones, ciertos detalles son sólo accesibles a algunos grupos de expertos, tales como que "oro" refiere a aquella sustancia que posee la propiedad de tener el número atómico 79 o que "agua" refiere a aquella sustancia que posee la propiedad de estar compuesta por dos moléculas de hidrógeno y una de oxígeno; sin embargo, el hecho de que el hablante común no pueda efectuar tales distinciones no implica que sea incapaz de usar correctamente los términos correspondientes: por el contrario, en virtud de pertenecer a la misma comunidad lingüística que los expertos, puede apoyarse en el conocimiento que éstos poseen y tomar prestada la referencia de los términos en cuestión.

De acuerdo con esto, la hipótesis de la división del trabajo lingüístico da lugar a una explicación puramente histórico-causal del mecanismo de transmisión de la referencia. Tenemos capacidad para hablar de estructuras químicas, electrones y genes, a pesar de que ignoremos todo (o estemos totalmente equivocados) acerca de tales cosas: la capacidad en cuestión no se basa entonces en el conocimiento individual sino en nuestros vínculos histórico-causales con ciertos grupos de expertos en esos temas. (Esto último podría perfectamente aplicarse a los nombres: una persona tiene la capacidad de referirse exitosamente a un individuo determinado mediante un nombre sólo en virtud de su inserción en una cadena histórico-causal

que desemboca, en última instancia, en el individuo en cuestión.) Esto es, en mi opinión, lo que quiere decir Putnam cuando afirma que no es el estado psicológico individual sino el estado sociolingüístico del cuerpo lingüístico colectivo lo que determina la referencia de los términos generales. En ningún caso, el concepto que cada individuo asocia con un cierto término determina la referencia del mismo.

En mi opinión, es posible diferenciar dos puntos distintos: por un lado, Putnam destaca la importancia del mundo mismo, independientemente del conocimiento que podamos tener de él, en la constitución del significado; por otro lado, destaca que el único tipo de conocimiento semánticamente pertinente es el conocimiento entendido como una dimensión social, el cual es perfectamente compatible con la más absoluta ignorancia individual. De acuerdo con esto, podría decirse que las teorías descripcionales cometen dos pecados: por un lado, ignoran la indicabilidad del lenguaje, es decir, el rasgo del lenguaje por el cual éste nos trae o nos conecta directamente con el mundo (defecto antes caracterizado como incompletitud de las teorías descripcionales); por otro lado, ignoran la dimensión social del conocimiento.

Cabe destacar en este punto que, si bien Putnam, mediante la propuesta de la hipótesis de la división del trabajo lingüístico, contribuye a iluminar de manera significativa el mecanismo (histórico-social) de transmisión de la referencia, no ocurre lo mismo con el de fijación de la misma. En este sentido, su teoría no involucra ninguna ventaja con respecto a la de Kripke: en la medida en que se comprometen con la afirmación de que la referencia es constituida exclusivamente por la relación causal, ambas son igualmente afectadas por el antes mencionado problema del respecto.

1.1.2. Una nueva dimensión epistémica: el estereotipo

La propuesta por parte de Putnam de un modelo causal muy similar al defendido por Kripke no debe conducir a pensar que Putnam pretende agotar la explicación del significado de los términos de clases naturales por medio de una teoría causal de la referencia directa -a la

manera de Kripke. Esto es, sin duda, un error: Putnam considera que, a fin de obtener una explicación completa del significado, una teoría causal de la referencia directa como la esbozada por Kripke debe ser complementada mediante una teoría del estereotipo; podría decirse entonces que Putnam propone una explicación puramente causal de la referencia pero no del significado de los términos generales.

¿Qué es un estereotipo? La respuesta a esta pregunta requiere un pequeño rodeo. Si bien la competencia lingüística con un término general no exige poseer un conocimiento identificador del referente, sí exige saber algo acerca de éste: de lo contrario, se considera que el hablante no ha adquirido el término en cuestión. Esto que el hablante competente está obligado a saber acerca de la referencia es una idea convencional, "frecuentemente maliciosa y harto imprecisa de cómo parece ser, es o se comporta" el objeto referido: un estereotipo. El estereotipo es, por tanto, una entidad claramente epistémica, constituida por lo que el hablante sabe acerca del objeto y puede expresar por medio de un conjunto impreciso de descripciones.

Ahora bien, el hecho de que una determinada descripción se incluya en el estereotipo asociado con un término general determinado no implica que sea verdadero ni mucho menos analíticamente verdadero que todas las muestras o ejemplares referidos por el término en cuestión tengan el rasgo mencionado por la descripción (ni siquiera que todas las muestras o ejemplares normales lo tengan). A modo de ejemplo, "un animal con rayas negras transversales y panza blanca" es parte del estereotipo de "tigre", por lo que, para adquirir el término "tigre", es obligatorio poseer la información de que los tigres estereotípicos son animales con rayas negras transversales y panza blanca; pero, dado que la obligatoriedad lingüística no implica verdad ni mucho menos analiticidad, la oración

> Los tigres son animales con rayas negras transversales y panza blanca

no sólo puede ser falsa sino que de hecho lo es. En general, el conocimiento de los rasgos seleccionados por el estereotipo no da lugar,

por tanto, a oraciones analíticas, necesarias y *a priori*, puesto que ni siquiera da lugar por lo general a oraciones verdaderas. Sin embargo, distintas oraciones tienen distintos grados de revisabilidad: la oración antes mencionada es más revisable que la siguiente:

Los tigres son animales

Esta última, si bien no absolutamente irrevisable, es más difícil de revisar que otras: la razón de ello es que predica un rasgo central o indicador de categoría, por lo que constituye, en la terminología de Putnam, un marcador semántico.

Por consiguiente, si bien Putnam reintroduce una dimensión epistémica en el significado, no se trata del conocimiento identificador exigido por las teorías descripcionales: el estereotipo de un término general de clase natural involucra el conocimiento no de condiciones necesarias y suficientes para la aplicación del término sino de un conjunto de características convencionales e imprecisas, a veces incluso engañosas, que el hablante competente asocia comúnmente con el término en cuestión. Este tipo de conocimiento da lugar no ya a una definición del término sino a una caracterización vaga e imprecisa del mismo.

Estamos ahora en condiciones de ejemplificar, para el caso del término "tigre" lo que Putnam denomina "forma normal de descripción del significado": se trata de un vector que reúne los siguientes componentes

(i) marcadores sintácticos: "nombre concreto"
(ii) marcadores semánticos: "animal"
(iii) descripción de los rasgos adicionales del estereotipo: "poseedor de rayas negras transversales"
(iv) descripción de la extensión: "Felis Tigris"

Los puntos (i), (ii) y (iii) hacen referencia al estereotipo, es decir, al conocimiento que determina la competencia del hablante individual; (iv), por el contrario, indica precisamente lo que trasciende la

competencia individual y es, por el contrario, aportado por el entorno. De este modo, el significado es fundamentalmente un par ordenado que contiene el estereotipo y la extensión o, en los términos que venimos utilizando, la referencia.

Ahora bien, vale la pena destacar que estos dos factores son, como se enunció más arriba, conceptualmente distintos y mutuamente independientes. A diferencia del sentido fregueano, el estereotipo no cumple la función semántica de determinar la referencia. En términos del ejemplo anterior, el hablante común asocia "rayado" con "tigre"; sin embargo, "tigre" refiere a la clase de los tigres en virtud de una relación causal directa con sus distintos ejemplares y no en virtud de ser la clase de objetos seleccionada por (entre otras descripciones) "rayado", dado que de hecho no lo es; en general, dado que las descripciones asociadas con los términos generales por el hablante común son por lo general falsas respecto de los objetos referidos, no puede apelarse a ellas como mecanismo de determinación de la referencia y la verdad. Sin embargo, en la medida en que hay razones independientes para considerar que constituyen una parte del significado (razones que para Putnam están, como hemos visto, en parte relacionadas con las condiciones de aprendizaje de los términos generales y, como se verá en el apartado siguiente, en parte, relacionadas con la necesidad de incorporar propiedades semánticas explicativas de la conducta), ha de considerarse que se trata de una parte radicalmente distinta e independiente de la dimensión referencial o representacional. De este modo, el *slogan* de Putnam debería ser no el antes antes citado –"meanings just ain't in the head!"- sino, más precisamente, este otro: "meanings ain't just in the head!", es decir, los significados no están *sólo* en la cabeza.[69]

De ahí que, como también se mencionó, Putnam considere que la explicación del significado requiere dos teorías distintas: por un lado, una teoría sociolingüística que se ocupa del problema de explicar los mecanismos de determinación de la referencia (en términos causales y con la ayuda de la antes mencionada hipótesis de la división

69. Debo la propuesta de modificación del *slogan* a Richard Hanley.

del trabajo lingüístico); por otro lado, una teoría psicolingüística que se dedica al problema de describir qué es lo que determina la competencia individual. La teoría de Putnam constituye por consiguiente la primera teoría semántica del doble factor: si bien, por un lado, Putnam se inscribe en la línea de Kripke para reivindicar la importancia del mundo externo en la determinación de la referencia, por otro, reintroduce un factor de índole descriptiva, y, por tanto, una dimensión epistémica en el significado. Esto ultimo confiere a su teoría la clara ventaja, por sobre la teoría de Kripke, de ofrecer una solución a lo que antes denominamos "el problema del sentido" y, consiguientemente, a los problemas semánticos tradicionales que resurgían, como espectros, ante la nueva propuesta causal.

Sus actuales sucesoras reemplazan la noción de estereotipo por la de rol conceptual.[70] Pero se trata esencialmente del mismo tipo de teorías: teorías que intentan conciliar el modelo causal inaugurado por Kripke -que enfatiza la influencia del entorno físico e histórico-social en la constitución del significado- con la tradición cartesiana –que atribuye un rol al conocimiento otorgado por las descripciones asociadas-, de la que el mismo Frege estaba profundamente imbuido. Véanse al respecto las siguientes palabras de Block:

> *La idea de la versión del doble factor es que el significado tiene dos componentes, un componente constituido por el rol conceptual que está enteramente 'en la cabeza' y un componente externo que tiene que ver con las relaciones entre las representaciones en la cabeza (con sus roles conceptuales internos) y los referentes y/o condiciones de verdad de esas representaciones en el mundo.*[71]

Ahora bien, en mi opinión, sin embargo, hay razones para rechazar este tipo de teorías; en otras palabras, si bien las teorías del doble

70. Para un ejemplo paradigmático, véase la teoría de Block presentada en Block 1986. Véanse también McGinn 1982, Loar 1982 y Field 1977. Esta última versión es especialmente compleja puesto que define la noción de rol conceptual en términos de la noción de función de probabilidad subjetiva.
71. Block 1986: p. 627.

factor poseen recursos –en particular, las nociones de estereotipo y rol conceptual- que permiten evitar ciertos problemas, lo hacen a un costo muy alto. En otras palabras, considero que las nociones de estereotipo y rol conceptual no permiten construir una explicación adecuada del funcionamiento semántico del lenguaje natural. Pero antes de explicitar mis razones en favor de esta tesis, creo oportuno describir el panorama teórico más amplio en el que dichas teorías se insertan, de modo tal que resulte patente la tensión interna que las caracteriza. En ello está centrado el siguiente apartado.

1.2. Dos tradiciones para las teorías del doble factor

Como se sugirió anteriormente, es posible considerar que las teorías del significado del doble factor se recortan sobre el fondo de dos grandes tradiciones: por un lado, la gnoseología cartesiana, según la cual sólo lo que ocurre dentro de la cabeza, esto es, las entidades mentales denominadas "ideas", son objeto de conocimiento directo, mientras que la existencia del mundo externo es inferida a partir de aquéllas;[72] y, por otro lado, el reconocimiento contemporáneo, por parte tanto de filósofos como de psicólogos, del rol del entorno en la determinación del conocimiento, el significado y la conducta. En lo que sigue, me interesa enfatizar las diferencias entre las concepciones y teorías originadas por cada una de ellas, tanto en el ámbito estrictamente semántico como en el ámbito de la filosofía de la mente. Mi objetivo es, en última instancia, desentrañar la naturaleza dual que caracteriza a las teorías del doble factor.

72. Recuérdese la tradición moderna, encarnada en los realistas representativos, esto es, fundamentalmente Descartes, Locke y Hume: desde esta perspectiva, las ideas son entidades mentales, subjetivas y directamente accesibles a la conciencia (por introspección), que constituyen los significados de las palabras del lenguaje natural.

1.2.1. Sentidos, roles conceptuales y referentes

La concepción fregueana del sentido expuesta en el capítulo anterior constituye un ejemplo de la influencia, en semántica, de la primera tradición mencionada, es decir, la de la gnoseología cartesiana. Es cierto que Frege se propuso destacar el carácter público de los sentidos e insistió, por tanto, en que se trataba no de entidades mentales sino de entidades abstractas. Pero, el problema es, como señala Putnam, que captar una entidad abstracta sigue siendo un acto mental o psicológico; de esta manera, la teoría del sentido de Frege se halla asociada con una teoría mentalista, y, en esa medida, de influencia cartesiana, de la comprensión del sentido.[73] Alternativamente, si se deja de lado la sugerencia fregueana de interpretar a los sentidos como entidades abstractas y se los considera, como se sugirió en el capítulo anterior, equivalentes a conceptos o representaciones conceptuales, a la manera de Salmon, los sentidos se revelan claramente como entidades mentales. De acuerdo con esto, los sentidos, en tanto modos de presentación de los objetos referidos, constituyen los diferentes estados mentales en los que aquéllos son aprehendidos. Como vimos, la noción fregueana es la piedra fundacional de la concepción descripcional de la referencia, según la cual el referente de un término está determinado o bien por una sola descripción o bien por un cierto subconjunto ponderado del conjunto de descripciones asociadas con el mismo por los hablantes competentes, es decir, en ambos casos, por entidades que sirven para caracterizar estados mentales que median la relación con el mundo.[74]

73. Véase Putnam 1975: p. 222 (texto citado anteriormente en la nota 27).
74. Cabe aclarar que, como señala Putnam en el texto largo citado anteriormente, afirmar que los sentidos son entidades mentales no equivale a afirmar que son entidades *individuales*, esto es, que varían de individuo a individuo (e incluso en el mismo individuo a lo largo del tiempo). El mismo tipo de entidad mental puede instanciarse en dos mentes distintas, como muestra el hecho de que individuos distintos puedan tener los mismos conceptos. Sin embargo, una vez que se deja de lado el supuesto carácter abstracto de los sentidos fregueanos, no está del todo claro qué es lo que garantiza su carácter *intersubjetivo*.

Mientras Kripke y Putnam gestaban el nuevo modelo causal, otros filósofos, influidos por el desarrollo de las teorías funcionalistas de la mente, la psicología cognitiva y la teoría chomskyana de la sintaxis de los lenguajes naturales, propusieron nuevas teorías semánticas de fuerte inspiración mentalista o racionalista, aun más que las teorías descripcionales antes mencionadas: las denominadas "semánticas de roles conceptuales".[75] De acuerdo con éstas, el significado de un término está dado por el concepto para cuya expresión el término es usado, y el significado de aquél está a su vez determinado por el rol funcional que cumple en la psicología de una persona. El rol en cuestión depende de sus posibles combinaciones con otros conceptos, así como con inputs sensoriales y outputs conductuales. Según Harman,

> *"Rol funcional" incluye cualquier rol especial que un concepto pueda desempeñar en la percepción y en la inferencia o razonamiento, incluyendo el razonamiento práctico que conduce a la acción.*[76]

En otras palabras, el rol funcional de un concepto es su rol inferencial, constituido por las llamadas "propiedades inferenciales" del concepto en cuestión, es decir, aquellas propiedades determinadas por las relaciones inferenciales entre las oraciones en las que puede figurar el concepto -las cuales se dan entre las oraciones en virtud de la presencia de ese concepto. Nuevamente, en palabras de Block:

> *(...) el rol conceptual, depende del rol causal de la expresión en el razonamiento y la deliberación y, en general, en la manera en que la expresión se combina e interactúa con otras expresiones de manera tal de mediar entre los inputs sensoriales y los outputs conductuales. Un componente crucial del rol conceptual de una oración depende de su participación en inferencias inductivas y deductivas. El rol conceptual de una palabra depende de su contribución al rol de las oraciones.*[77]

75. Véase especialmente Fodor 1981 y Harman 1982.
76. Harman 1982: p. 242.
77. Block 1986: p. 628.

Cabe aclarar entonces que las propiedades inferenciales dependen no sólo de relaciones de inferencia deductiva entre las oraciones sino también de asociaciones subjetivas e individuales, que varían de acuerdo con los rasgos psicológicos de cada uno. Por consiguiente, el rol conceptual de una expresión es una entidad no sólo mental sino también individual. De este modo, la noción de rol conceptual tiene una clara dimensión epistémica: representa un conjunto de propiedades claramente accesibles al sujeto. Intuitivamente, se tiende a establecer una relación causal entre la posesión (o la carencia) del conocimiento involucrado por los roles conceptuales por parte de un sujeto determinado y su modo de actuar. De acuerdo con esto, el propósito fundamental de la adscripción de significados es, para estas teorías, la explicación de la conducta.

Tal vez sea útil explicitar en este punto de qué manera la adscripción de roles conceptuales hace posible la explicación de la conducta. Para retomar el ejemplo presentado al final del capítulo anterior, la decisión por parte de Pedro de no ir al concierto puede ser explicada en términos de su creencia de que tocará el primer violinista del cuarteto local, es decir, mediante la adscripción a Pedro del concepto EL PRIMER VIOLINISTA DEL CUARTETO LOCAL; si su actitud hubiera sido distinta (si Pedro hubiera ido al concierto), tal vez se habría optado, para explicarla, por la adscripción de otro concepto, tal como EL RECIENTE GANADOR DEL PREMIO BACH. EL PRIMER VIOLINISTA DEL CUARTETO LOCAL y EL RECIENTE GANADOR DEL PREMIO BACH juegan roles claramente distintos en la psicología de Pedro o, en otros términos, tienen distintos roles inferenciales.

A los fines de la explicación de la conducta, parece entonces necesario concentrarse en el rol inferencial que los diferentes conceptos, asociados con distintas expresiones, juegan en la psicología del agente. Como en el caso de la solución de los problemas semánticos tradicionales antes mencionados, se trata de un propósito semántico que parece requerir en la mayor parte de los casos la adscripción no de referentes sino de significados que poseen, en virtud de sus componentes epistémicos, un mayor grado de discriminación. Es precisamente la base epistémica de este rasgo lo que acerca los roles

conceptuales a los sentidos y cúmulos de descripciones de las teorías descripcionales clásica y refinada respectivamente. De este modo, es posible afirmar que las semánticas de roles conceptuales, al igual que las teorías descripcionales, son herederas contemporáneas de la tradición cartesiana en semántica. Más aún, en la medida en que las semánticas de roles conceptuales, a diferencia de las teorías descripcionales, eliminan por completo la dimensión referencial, pueden ser consideradas el ejemplo más conspicuo de la mencionada tradición.[78] Tal vez valga la pena destacar que estas teorías son, en virtud de este rasgo, teorías no representacionalistas o antirrepresentacionalistas del significado: de acuerdo con ellas, ninguna de las propiedades inferenciales que lo constituyen son propiedades que contribuyan a determinar la referencia, puesto que no hay tal cosa; en la medida en que no postulan componentes representacionales del significado, no puede considerarse, a diferencia también de lo que ocurre con las teorías descripcionales, que den lugar a instancias de la concepción correspondentista de la verdad. En este aspecto, se asemejan a las teorías del significado de tipo verificacionista, esto es, a grandes rasgos, aquéllas que identifican el significado con condiciones de verificación o de uso, es decir, condiciones subjetivas.

La segunda tradición mencionada anteriormente podría llamarse "naturalismo". En la medida en que reconocen el rol crucial del entorno en la determinación de la conducta, los pragmatistas clásicos, Peirce, Dewey y James, pueden ser considerados representantes paradigmáticos de este marco teórico alternativo. Pero son Kripke y (aunque parcialmente) el mismo Putnam quienes abren al naturalismo las puertas de la semántica, al señalar la necesidad de considerar al referente (o referencia o extensión) como componente directo del significado de los términos. Como se enfatizó oportunamente, el experimento mental de la Tierra Gemela intenta mostrar que la identificación del rol conceptual asociado con

78. Para estas teorías, a diferencia de lo que ocurre con las teorías descripcionales, la dimensión epistémica no tiene la función semántica de determinar la referencia puesto que agota en sí misma el concepto de significado.

una expresión (o lo que él llama, como vimos, "el estereotipo asociado con una expresión") no es suficiente para determinar su referente: es necesario tener en cuenta la relación directa entre la expresión y el mundo externo. Manteniendo el ejemplo de Putnam, la relación con H_2O y la misma H_2O son tan constitutivas del significado de "agua" como el conjunto de sus relaciones inferenciales con "líquido", "transparente", "refrescante", etc y sus relaciones causales con outputs conductuales como el calmar la sed. De acuerdo con esto, dos hablantes pueden asignar el mismo rol conceptual a una expresión dada y, sin embargo, la referencia de la misma en el idiolecto de uno de ellos puede diferir de su referencia en el idiolecto del otro -como es el caso de Oscar 1 y Oscar 2 mencionado en el experimento.

Las expresiones indicadoras proporcionan un ejemplo claro del rol directo desempeñado por la referencia en la constitución del significado. A modo de ejemplo, la oración

Tengo que trabajar

tiene el mismo rol conceptual cuando es emitida por Ana y por María: está relacionada con el mismo tipo de oraciones, tales como "Necesito dinero" o "Estoy muy ocupada", y produce el mismo tipo de conductas. Sin embargo, es claro que sus distintas emisiones no tienen el mismo significado, puesto que la primera refiere a Ana mientras que la segunda refiere a María.[79]

Más aún, como se vio en parte a propósito de la teoría de Kripke, las relaciones de referencia directa son necesarias para resolver problemas de ignorancia (casos en los que la conexión concepto-objeto no está mediatizada por ninguna descripción accesible al sujeto) y representaciones erróneas (casos en los que un concepto está conectado con el objeto equivocado).

79. Véase al respecto la teoría de la referencia directa desarrollada para los demostrativos por Kaplan en Kaplan 1978 y 1989.

Además, como también se mencionó anteriormente, se suele considerar que la adscripción de relaciones referenciales directas responde a un propósito semántico distinto del de la explicación de la conducta: el incremento de nuestro conocimiento acerca del mundo, que nos conduce del plano del pensamiento al plano de las condiciones de verdad. Aun en los casos en los que se adscriben pensamientos finamente discriminados, existe el deseo de establecer si son verdaderos o falsos, mostrando de esta manera un interés que va más allá del nivel de los roles conceptuales, un interés por saber cómo el mundo realmente es independientemente de lo que creamos o sepamos acerca de él.

1.2.2. Internalismo y externalismo

Es pertinente señalar que la crítica de Putnam a las teorías descripcionales y, en general, a las teorías mentalistas, está estrechamente relacionada y, más específicamente, constituye de hecho el detonante principal de un famoso debate en filosofía de la mente: el debate internalismo-externalismo, cuyo núcleo es semántico pero tiene importantes derivaciones metafísicas.[80] A fin de hacer explícito el núcleo en cuestión, considero conveniente explicitar qué se entiende por "estado mental", es decir, cómo es posible definir el concepto clave de la filosofía de la mente. En general, se considera que los estados mentales se individualizan en función de: (i)sus tipos: pueden ser puramente intencionales o actitudes proposicionales, tales como las creencias, los deseos y las intenciones, mixtos, esto es, intencionales y fenoménicos o cualitativos, tales como las percepciones, las imágenes y las cadenas de pensamientos, o puramente cualitativos, tales como los dolores y algunas otras sensaciones corporales; y (ii) sus contenidos. Ahora bien, según la concepción relacional mencionada en el capítulo anterior, el contenido intencional de una actitud proposicional (que es el tipo de estado mental que nos concierne en este trabajo y del

80. Fodor y Burge son algunos de sus protagonistas más conspicuos. Véanse, por ejemplo, Fodor 1991 y Burge 1979, 1986 y 1988.

cual puede decirse a su vez que comprende distintos modos, como las creencias, los deseos, etc.) está dado por el significado de una oración del lenguaje con la cual el sujeto de la actitud proposicional tiene un cierto tipo de relación, a saber, la oración denominada "adscripta" –la cual, como se mencionó anteriormente, puede considerarse que pertenece o bien al lenguaje natural o bien al lenguaje mental; a modo de ejemplo, el contenido intencional de la creencia de que la tarde es lenta puede ser especificado en términos del significado de la oración (sea ésta del lenguaje natural o del lenguaje mental) "La tarde es lenta".[81] Este último rasgo hace, por tanto, patente el núcleo semántico del debate.

Ahora bien, por un lado, como señala el propio Putnam, las teorías descripcionales y las semánticas de roles conceptuales, recortadas sobre el fondo de la tradición cartesiana, implican una concepción de los estados mentales que presupone lo que denomina "solipsismo metodológico":

> *[...] la asunción de que ningún estado psicológico, propiamente llamado así, presupone la existencia de ningún otro individuo excepto el sujeto al cual el estado es adscripto.*[82]

La aceptación del solipsismo metodológico equivale a la adopción de un criterio intrínseco, individualista o no relacional para la individualización de los estados mentales. Los contenidos de tales estados se conocen en general como "contenidos estrictos" ("*narrow*

81. En un sentido más amplio de "contenido", puede decirse que la oración en cuestión también tiene *contenido,* así como la oración (principal y subordinada) del lenguaje natural en el que se realiza la adscripción. En definitiva, se puede y suele decir que tienen contenido tanto las actitudes proposicionales como las expresiones del lenguaje natural y del lenguaje mental; en estos últimos sentidos, "contenido" es sinónimo de "significado". Cabe aclarar, sin embargo, que es más común utilizar la expresión "contenido" en lugar de "significado" en relación con las expresiones del lenguaje mental, es decir, los pensamientos y los conceptos, que hacerlo en relación con las expresiones del lenguaje natural.
82. Putnam 1975: p. 220.

contents"). De este modo, los sentidos fregueanos y los roles conceptuales, en función de los cuales se determinan los contenidos en cuestión, pueden ser a su vez considerados significados estrictos (*narrow meanings*).

La cara metafísica del debate, anunciada más arriba, está dada por el hecho de que las teorías semánticas que defienden el solipsismo metodológico son teorías internalistas, es decir, teorías que respetan la superveniencia mente-cerebro. Por relación de superveniencia entre propiedades se entiende lo siguiente: una propiedad A superviene en una propiedad de base B si y sólo si A está constituida o compuesta por B sin ser idéntica a B, de modo tal que necesariamente si A cambia entonces B cambia -nótese que la tesis de la superveniencia involucra una afirmación modal. El internalismo (también llamado "individualismo") es la posición en filosofía de la mente según la cual, necesariamente, si dos individuos comparten todas sus características físicas (es decir, si son molecularmente idénticos) entonces comparten sus estados psicológicos (es decir, son psicológicamente idénticos); en términos de la noción de superveniencia antes mencionada, es la tesis de que las propiedades o estados mentales de un individuo supervienen en sus propiedades o estados físicos (más específicamente, cerebrales) -por lo que los primeros no pueden cambiar sin que cambien los segundos. Como puede verse, el internalismo involucra entonces la adopción de un criterio intrínseco o no relacional de individualización de los estados mentales, es decir, una individualización realizada en función de los significados estrictos.

Por otro lado, las consideraciones de los teóricos causales marcan una dirección opuesta a la antes mencionada: lejos de sostener el solipsismo metodológico, señalan la necesidad de adoptar un criterio relacional para la individualización de los estados mentales, es decir, un criterio que tome en cuenta la manera en que aquéllos se relacionan directamente con el mundo exterior a la mente. Los contenidos de tales estados se conocen en general como "contenidos amplios" ("*wide contents*"). De este modo, los referentes y las condiciones de verdad, en función de los cuales se determinan los contenidos en cuestión, pueden ser a su vez considerados significados amplios (*wide meanings*).

Como puede preverse, las teorías semánticas que rechazan el solipsismo metodológico, en particular, las teorías histórico-causales de Kripke y Putnam expuestas anteriormente, la teoría de Kaplan acerca de los demostrativos, la teoría de Burge, dan lugar a teorías externalistas de la mente, es decir, teorías que no respetan la superveniencia mente-cerebro. El externalismo es la posición en filosofía de la mente según la cual puede haber identidad física o molecular sin identidad psicológica; en otras palabras, las propiedades o estados mentales no supervienen en las propiedades o estados cerebrales -por lo que los primeros pueden cambiar sin que lo hagan los segundos.[83] El externalismo presupone entonces la adopción de un criterio relacional de individualización de los estados mentales, esto es, realizada en función de los contenidos amplios.

Ahora bien, las teorías del doble factor, como la ofrecida por Putnam, no pertenecen a ninguno de los dos bandos en disputa, puesto que surgen justamente del esfuerzo por conciliar las dos tradiciones mencionadas. Esto puede parecer increíble, dado que el nombre "Putnam" aparece mencionado repetidas veces en relación con el rechazo de los significados estrictos, el solipsismo metodológico y el internalismo, por un lado, y la defensa de los significados amplios y el externalismo, por otro. Sin embargo, no hay que olvidar que esto constituye sólo una parte de su teoría, la que se ha caracterizado como "teoría sociolingüística": es sólo esta parte de la teoría de Putnam la que puede considerarse que ejemplifica el modelo histórico-causal. Alternativamente, es posible considerar que Putnam propone no una sino dos teorías distintas, y que sólo una de ellas es de carácter externalista; la otra, la teoría del estereotipo, también denominada "teoría psicolingüística", es, por el contrario, de carácter claramente internalista. El carácter mixto de la posición de Putnam responde a lo dicho anteriormente: por un lado, Putnam niega que el significado de un término esté agotado por su sentido, como sostienen los defensores de las semánticas de roles conceptuales, o por su referente, como

83. Dado que la tesis de la superveniencia es una tesis modal, para rechazarla basta con señalar la *posibilidad* de un contraejemplo. Véase Sabatés 1993.

sostienen los partidarios de las teorías de la referencia directa; por otro lado, también niega que el sentido determine al referente, a la manera de los teóricos descripcionales. En su opinión, sentido o, con mayor propiedad, estereotipo y referente son dos aspectos radicalmente distintos e independientes que constituyen conjuntamente el significado. De ahí que la explicación del mismo requiera dos teorías radicalmente distintas e independientes. De este modo, Putnam, lejos de pertenecer a alguno de los dos bandos en disputa, representa de alguna manera a ambos -la restricción incluida está relacionada con el hecho de que la combinación de dos teorías da lugar a un enfoque mixto que en tanto tal es más amplio y complejo que los enfoques involucrados por cada una de las teorías independientemente consideradas. Por consiguiente, tal vez sea más adecuado afirmar que, si bien está influido por ambos, de alguna manera Putnam no representa a ninguno de los dos bandos en disputa.

Independientemente de cómo se quiera describir el lugar teórico que ocupa Putnam, hay dos puntos, también mencionados previamente, que quisiera enfatizar. En primer lugar, los dos aspectos reconocidos en el significado son considerados conceptualmente distintos y mutuamente independientes. Se rechaza entonces el presupuesto fregueano según el cual las propiedades inferenciales constituyen el significado sólo en la medida en que determinan la referencia -puesto que se considera que ninguna propiedad inferencial determina la referencia, si bien todas ellas constituyen el significado. En segundo lugar, cada aspecto del significado cumple un propósito semántico distinto: mientras que la adscripción de estereotipos da lugar a una explicación adecuada de la conducta, la asignación de referentes y condiciones de verdad hace posible el aumento del conocimiento acerca del mundo -siempre que se considere a los hablantes lo suficientemente confiables. Son precisamente estos dos puntos los que encuentro cuestionables: a ellos está dirigida fundamentalmente la crítica desarrollada a continuación. Dado que los puntos en cuestión son comunes a todas las teorías del doble factor, la crítica no se aplica sólo a la teoría de Putnam sino a todas ellas -y hará referencia consecuentemente no a los términos particulares de la teoría de Putnam sino a los términos generales de cualquier teoría de su tipo.

1.3. Un análisis crítico

1.3.1. ¿Teoría descriptiva o teoría normativa?

Ante todo, cabe aclarar que en este apartado se presupone el punto de vista de la metodología propuesta al final del capítulo anterior, según la cual se parte del análisis de las adscripciones de actitudes proposicionales para establecer la naturaleza no del significado de aquéllas (es decir, el significado de las oraciones adscriptoras o significado de segundo nivel) sino del significado de lo adscripto por medio de ellas (es decir, el significado de las oraciones adscriptas o significado de primer nivel).[84] Ahora bien, creo que es importante tratar de establecer si los teóricos del doble factor consideran que están describiendo los significados que de hecho adscribimos o si se consideran a sí mismos revisionistas, es decir, si creen estar proponiendo los significados que deberíamos adscribir a la luz de los propósitos semánticos. En el primer caso, la teoría defendida sería descriptiva; en el segundo, en cambio, se trataría de una teoría normativa. La importancia de aclarar este punto reside, en mi opinión, en que de este modo se facilita la búsqueda y evaluación de los argumentos que podrían avalar a estas teorías.

La discusión subsiguiente se basa asimismo en el uso de la difundida distinción conceptual mencionada en el capítulo anterior: la distinción entre adscripciones transparentes y adscripciones opacas o, en otros términos, adscripciones *de re* y adscripciones *de dicto*. Cabe recordar que las adscripciones transparentes o *de re* involucran una relación referencial directa con un objeto específico; en otros términos, son adscripciones de referentes. Las adscripciones opacas o *de dicto*, en cambio, atribuyen una relación referencial con un objeto bajo un modo específico de presentación del objeto en cuestión; en otros términos, son adscripciones de modos de presentación

84. En términos de éste último puede especificarse, como se señaló en el apartado anterior, el *contenido* de las actitudes proposicionales correspondientes.

o sentidos. Ahora bien, éstas últimas también pueden ser entendidas como adscripciones de roles conceptuales: en general, se trata de adscripciones de significados estrictos, que no pueden ser concebidos en términos de los objetos referidos. Me referiré a estos significados estrictos, de una manera general, como "modos" o "sentidos".

1.3.1.1. El problema de la evidencia

Presupondré, para empezar, que la primera opción antes mecionada es la correcta, es decir, que los teóricos del doble factor están describiendo los significados que de hecho adscribimos: de acuerdo con esta presuposición, adscribimos tanto significados estrictos como amplios, y esta práctica usual debe entenderse como una adscripción simultánea de estereotipos o roles conceptuales, por un lado, y condiciones veritativas, por el otro. Según la ocasión, un factor adquiere más relevancia que el otro: a veces, estamos más interesados en el rol funcional de un concepto; otras, en el objeto referido. La tesis descriptiva afirma entonces que adscribimos simultáneamente dos factores de distinta naturaleza (requisito de simultaneidad). Si esta tesis es verdadera, deberá encontrarse evidencia en su favor. Pero, en mi opinión, tal evidencia no existe. El punto es entonces que no queda claro de qué manera el requisito de simultaneidad puede ajustarse a nuestra práctica lingüística habitual.

Una adscripción transparente involucra una relación referencial directa con un objeto específico, sin importar bajo qué descripción se lo presenta o con qué conceptos se lo asocie. A modo de ejemplo, imaginemos la siguiente situación: el doctor Freudenthal es un reconocido psicoanalista que atiende a una paciente llamada "Lucila Corral". La secretaria del médico, íntima amiga de la paciente, escucha una conversación de éste con un colega en la que afirma que la paciente Corral está loca. A continuación, al encontrarse en la calle con una amiga común, la secretaria comenta, haciendo uso de un sobrenombre de Lucila sólo utilizado entre familiares y amigos y, por tanto, ignorado por el médico:

El doctor Freudenthal cree que Fafa está loca

Al emitir la oración, la secretaria atribuye a Freudenthal una creencia acerca de una persona, Lucila Corral, independientemente de cómo esa persona pueda ser, y de hecho sea, descripta por el médico. Por supuesto, se refiere a esa persona de una manera (mediante el sobrenombre "Fafa") que puede ser comprendida por su interlocutora, pero esto no nos interesa a los fines presentes: el punto es que no atribuye al doctor Freudenthal una creencia que involucra a "Fafa" (ni a ninguna otra descripción en particular) sino una creencia que involucra a Fafa, cualquiera sea la manera en que se la describa.

Por otro lado, una adscripción opaca involucra una relación referencial con un objeto bajo un modo específico. Por ejemplo, supongamos que siempre que Lucila Corral concurre a sus sesiones de terapia, se pone un velo negro sobre la cara; el encargado del edificio, que la ve entrar todos los días, la considera una persona muy extraña; sin embargo, no opina lo mismo de la mujer que encuentra todas las mañanas en el mercado, la cual, insospechadamente para él, no es otra que Lucila Corral. La adscripción

El encargado del edificio cree que la mujer del velo negro está loca

atribuye al encargado una creencia acerca de una persona, Lucila Corral, descripta de cierta manera, como la mujer del velo negro. Por supuesto, la descripción en cuestión es la descripción de un individuo; pero el punto es, sin embargo, que se atribuye a alguien no una creencia que involucra a Lucila Corral bajo cualquier descripción que se haga de ella -"la mujer que hace las compras con un canasto rojo", "Fafa", etc.- sino una creencia que involucra a esa persona descripta de cierta manera específica, esto es, como la mujer que usa un velo negro.

De acuerdo con lo anterior, parecería que, en cada ocasión particular, el significado adscripto es o bien un referente específico o bien un modo específico de presentación de un referente, en términos generales, un modo o sentido específico; por lo tanto, no hay casos, como sostienen las teorías del doble factor, en los que se adscriban simultáneamente ambos factores. Más que un aspecto distinto del significado adscripto, cada factor parece constituir un significado alternativo.

El requisito de simultaneidad parece involucrar una confusión entre el significado de la adscripción y el significado adscripto mediante esa adscripción. Una cierta adscripción de significado puede ser ambigua, es decir, puede no resultar claro qué significado ha sido adscripto; pero esto no implica que la oración adscripta sea ambigua, es decir, que se haya adscripto más de un significado. El carácter ambiguo de las adscripciones de significado, que permiten tanto lecturas opacas como lecturas transparentes, no tiene porqué ser transmitido a las oraciones adscriptas. En términos del ejemplo anterior,

El doctor Freudenthal cree que Fafa está loca

puede resultar ambigua a quien ignore, por ejemplo, que el médico desconoce por completo el sobrenombre de su paciente: una adecuada información acerca del contexto apropiado le permitirá desambiguar la oración anterior. Pero la ambigüedad en cuestión no implica en modo alguno una simultánea adscripción inicial de dos significados distintos, a saber, un referente y un modo. Postular, como podría decirse que hacen los teóricos del doble factor, un modo referencial para adscripciones transparentes y un referente modal para opacas constituye una innecesaria multiplicación de entidades.

1.3.1.2. La cuestión normativa

Cabe ahora contemplar la segunda opción antes mencionada: considerar que las teorías del doble factor son revisionistas. Si éste fuera el caso, los teóricos del doble factor estarían afirmando que, más allá de los significados que de hecho adscribamos, deberíamos adscribir significados complejos, constituidos por dos factores independientes: el rol conceptual y el factor referencial. Estos son los significados que debemos adscribir, pues sólo este tipo de significados sirve a nuestros propósitos semánticos. De acuerdo con esto, se requiere un argumento en favor de la necesidad de la adscripción simultánea (requisito de simultaneidad) de dos factores independientes (requisito de independencia). Ahora bien, el único argumento que he encontrado es un argumento en favor del requisito de independencia. La mayoría de los teóricos del doble factor coinciden en afirmar que

cada factor responde a un propósito semántico distinto. En particular, señalan que la adscripción de roles conceptuales sirve exclusivamente para explicar la conducta, mientras que la adscripción de condiciones de verdad sirve exclusivamente para facilitar el proceso de aprendizaje acerca del mundo. Según ellos, es el hecho de que cada uno de los factores hace posible el cumplimiento de un propósito distinto lo que fundamenta su mutua independencia conceptual.[85]

En mi opinión, el argumento ofrecido no es bueno porque los propósitos semánticos mencionados pueden ser alcanzados mediante la adscripción de cualquiera de los dos factores en disputa, lo que muestra que no hay una unión conceptual entre ciertos propósitos y ciertos factores que fundamente la independencia mutua de éstos últimos. Por un lado, las adscripciones transparentes de significado, es decir, las adscripciones de relaciones referenciales con objetos específicos o de referentes, hacen posible la explicación de la conducta. Retomando el ejemplo anterior, imaginemos que Freudenthal recibe una carta amenazadora de Lucila Corral y, sin saber que se trata de ella, llama a la secretaria para comentarle el hecho con mucho miedo y preocupación. La secretaria reconoce la letra de Lucila. Cuando su amiga le pregunta por teléfono qué le ocurre al médico, a quien oyó quejarse con tono preocupado, la secretaria le contesta con la siguiente adscripción transparente:

El doctor Freudenthal cree que Fafa es peligrosa

Las adscripciones transparentes sirven como explicaciones de actitudes y conductas bajo el presupuesto de que se siguen del producto de buenas adscripciones opacas y ciertas oraciones de identidad. De este modo, la oración anterior explica la actitud del médico, la preocupación por Lucila, bajo la presuposición de que hay un término "a" tal que "a" = "Fafa", y, de manera opaca, Freudenthal está preocupado por a y Freudenthal cree que a es peligrosa.

85. Vale la pena destacar la diferencia entre esta tesis y la tesis defendida en el apartado anterior, según la cual el hecho de que en cada caso determinado uno de los factores haga posible el cumplimiento de un propósito determinado fundamenta la no simultaneidad de los factores en cuestión.

Por otro lado, si las adscripciones transparentes sirven para el propósito de aprender acerca del mundo, entonces también lo harán las opacas, dado que éstas últimas proporcionan aún más información acerca del mundo que las primeras: no sólo nos dicen acerca de qué objeto se tiene una creencia sino también bajo qué modo específico se la tiene. El doctor Freudenthal puede no ser confiable respecto de su juicio acerca de la salud mental de Lucila bajo ciertas descripciones, tales como "la mujer que hace las compras con un canasto rojo" o incluso "Fafa", pero sí puede serlo bajo otros, tales como "la paciente más difícil de la clínica neuropsiquiátrica". Luego, la forma opaca nos permite evaluar la probabilidad de verdad que tienen las creencias adscriptas. En términos de un nuevo ejemplo: supongamos que una maestra está preparando un examen consistente en una lista de oraciones que deberán ser clasificadas por los alumnos como verdaderas o falsas. La decisión de incluir

> Los cronistas de aquella época creyeron que Tirso de Molina había escrito la fraudulenta segunda versión del Quijote

en lugar de

> Los cronistas de aquella época creyeron que Fray Gabriel Téllez había escrito la fraudulenta segunda versión del Quijote

sin duda facilitará a los alumnos la elección de la respuesta correcta. Algunas maneras de representar el mundo contribuyen al proceso de aprendizaje más que otras.

En mi opinión, la razón por la que ambos propósitos semánticos pueden ser alcanzados mediante la adscripción de cualquiera de los dos factores reside en el hecho de que tales propósitos, a saber, la explicación de la conducta y el incremento de nuestro conocimiento acerca del mundo, lejos de ser mutuamente independientes, están estrechamente relacionados entre sí, hasta el punto de que en última instancia ambos son reductibles a uno solo: el incremento de nuestro conocimiento acerca del mundo. Lo que hace posible la explicación de la conducta humana es el conocimiento de hechos relativos a las mentes de las personas, más específicamente, a sus deseos,

creencias y demás actitudes proposicionales. Pero, a menos que creamos que los hechos subjetivos tienen rasgos ontológicos peculiares (los llamados "*qualia*") que los tornan inaccesibles al observador externo, los hechos en cuestión son parte del mundo que nos rodea y acceder a hechos tales es incrementar de alguna manera nuestro conocimiento acerca del mundo. Si esto es así, no parece haber razón alguna para creer que la explicación de la conducta y el incremento de nuestro conocimiento acerca del mundo hayan de requerir la adscripción de significados de naturaleza radicalmente distinta.

El punto anterior pone de manifiesto un rasgo negado explícitamente por los teóricos del doble factor: el hecho de que la relación referencial con un objeto o referencia es siempre importante, aun en los casos de adscripciones opacas. Para que una adscripción sea verdadera, tiene que ser correcta no sólo respecto del modo sino también respecto del objeto referido. A modo de ejemplo, supongamos que estoy justificada a afirmar

> Pedro cree que el próximo presidente de la Argentina va a modificar la economía del país

Supongamos, además, que el Pedro Gemelo tiene una creencia cuyo contenido intencional está en parte constituido por el mismo rol conceptual que la creencia atribuida a Pedro. Ahora bien, la atribución de tal creencia a Pedro no sería adecuada, puesto que lo involucraría en una relación referencial no con el próximo presidente y la Argentina sino con el próximo presidente Gemelo y la Argentina Gemela, acerca de los cuales Pedro no tiene creencia alguna. Esto parece sugerir la idea de que los modos y los referentes no son mutuamente independientes de la manera en que sostienen los teóricos del doble factor o, en otras palabras, que la naturaleza de modos y referentes no es, como ellos sostienen, radicalmente distinta.

Si esto es así, el único argumento en favor del enfoque revisionista es claramente débil: fracasa en establecer la necesidad de adscribir significados constituidos por dos partes independientes y radicalmente distintas. Dado que los roles conceptuales son

por definición independientes de los referentes, el fracaso del argumento parece indicar que los modos no son roles conceptuales.

1.3.2. El holismo

Ante todo, cabe aclarar que el sentido de "holismo" al que me referiré en este apartado no es el mencionado en la introducción de este trabajo, según el cual una teoría holística del significado es una teoría que prescinde del concepto empírico de referencia; en este sentido, ninguna de las teorías analizadas en este trabajo es o puede ser considerada una teoría semántica holística. Lejos de ello, el sentido aquí aludido es el presente en la polémica de la semántica contemporánea que enfrenta dicha posición al localismo. Este nuevo concepto de holismo puede ser definido de distintas maneras: en términos de propiedades inferenciales, en términos de creencias asociadas y a la manera de Fodor & LePore.[86] De acuerdo con la primera definición, el holismo es la postura según la cual el significado de una representación está constituido por todas sus propiedades inferenciales.[87] En el caso mencionado en segundo término, el holismo puede caracterizarse como la tesis según la cual todas nuestras creencias asociadas con una representación determinada forman parte del significado de la misma; el punto es que, dado que todas las creencias están inferencialmente relacionadas entre sí, es el sistema total de creencias lo que, para el holista, determina el significado de cada representación del lenguaje.[88] Finalmente, según los autores mencionados, el holismo acerca de las propiedades semánticas es definido como la tesis según la

86. Véase Fodor & Lepore 1992, especialmente la introducción.
87. Cabe aclarar que éstas están constituidas por las relaciones que una representación tiene con otras en virtud de las relaciones inferenciales que las oraciones en las que aparece la primera tienen con las oraciones en las que aparecen las otras -en virtud de la presencia de la representación en cuestión.
88. Es claro que no se trata *sólo* de las creencias verdaderas sino de *todas* las creencias. Por ejemplo, un niño podría creer que un picaflor es lo que su madre llama "picaflor"; la creencia en cuestión podría ser falsa (su madre bien podría llamar "picaflor" a los gorriones) y, sin embargo, para el holista, no por ello dejaría de constituir el significado de "picaflor".

cual una representación no puede tener significado a menos que muchas otras también lo tengan. Si bien estas formulaciones no son estrictamente equivalentes entre sí, todas ellas ponen de manifiesto el rasgo más característico de cada una de las posturas en disputa: la interdependencia semántica de las representaciones en el caso del holismo y la total independencia semántica de las mismas en el extremo opuesto, esto es, el atomismo o localismo atomista.

Ahora bien, es claro que las teorías comprometidas con la noción de rol conceptual, tales como las semánticas de roles conceptuales y las del doble factor, son en esa misma medida teorías semánticas holísticas. Más específicamente, si uno considera que el significado se identifica (aunque sea parcialmente) con el rol conceptual de un término, el rol que éste juega en la psicología de un individuo en virtud de sus propiedades inferenciales, entonces no podrá evitar concluir que todas las propiedades inferenciales constituyen el significado del término en cuestión. Como se señaló anteriormente, el rol conceptual de un término es su rol inferencial, pero no en el sentido lógico del término sino en un sentido más laxo según el cual las inferencias inductivas, las materiales y las asociaciones subjetivas contribuyen a determinar el rol en cuestión (y como consecuencia, el significado). Una concepción semejante implica el holismo porque no es posible apelar a criterio alguno que permita imponer restricciones al conjunto de inferencias y asociaciones admisibles; en otras palabras, no es posible proponer un criterio que permita distinguir, entre todas las propiedades inferenciales de un término dado, a un presunto subconjunto constitutivo del significado en forma exclusiva: el rol conceptual de un término es, por tanto, su rol inferencial total. En términos de Block:

> *El rol conceptual es el rol causal total, descripto de manera abstracta. (...) El rol conceptual deja de lado todas las relaciones causales excepto aquéllas que median en las inferencias, inductivas o deductivas, procedimientos de decisión y demás inferencias similares.*[89]

89. Block 1986: p. 628.

De ahí que pueda decirse que tanto las semánticas de roles conceptuales como las teorías del doble factor (entre otras) adhieren implícitamente al siguiente argumento en favor del holismo:

(i) Algunas propiedades inferenciales de una expresión constituyen su significado.

(ii) Si algunas propiedades inferenciales constituyen el significado, todas lo hacen.

Luego, (iii) Todas las propiedades inferenciales de una expresión constituyen su significado.

Es claro que hay distintas maneras de rechazar la premisa (ii), que constituye la premisa clave de este argumento. Sin embargo, el punto es que ninguna de ellas es accesible al tipo de teorías en estudio. A modo de ejemplo, se ha alegado que un criterio para distinguir a las propiedades inferenciales constitutivas del significado de aquéllas que no lo son es el provisto por la distinción entre lo analítico y lo sintético: de acuerdo con esto, las inferencias analíticas darían lugar a propiedades inferenciales constitutivas mientras que las inferencias sintéticas no lo harían -de este modo, podría decirse que "hombre" es una propiedad inferencial de "soltero" mientras que "histérico" no lo es, puesto que la inferencia que relaciona a

Juan es soltero

con

Juan es hombre

es analítica (esto es, el condicional que tiene por antecedente a la primera oración y por consecuente ala segunda es verdadero en virtud de su significado), mientras que la que lo relaciona con

Juan es histérico

es sintética (es decir, el condicional correspondiente es verdadero en virtud de un hecho). Sin embargo, por un lado, cabe tomar en cuenta la demoledora crítica desarrollada por Quine respecto de la distinción en cuestión.[90] Por lo demás, es preciso consignar que la posibilidad de apelar, con los fines antes mencionados, a la distinción analítico-sintético no está realmente abierta para los teóricos del doble factor.[91] Si se considera, a diferencia de Frege, que el sentido y la referencia/verdad son dos dimensiones completamente independientes del significado, las propiedades inferenciales que constituyen el significado no determinan la referencia. Si esto es así, los vínculos determinantes del significado (es decir, las relaciones inferenciales entre oraciones) no sólo no implican la verdad analítica de ninguna oración sino que son incluso compatibles con su falsedad. Por consiguiente, no es posible apelar a la distinción entre oraciones analíticamente verdaderas y oraciones sintéticamente verdaderas para a su vez distinguir entre propiedades inferenciales constitutivas y no constitutivas del significado. En términos del ejemplo anterior, no es posible sostener que "hombre", a diferencia de "histérico", forma parte del significado de "soltero" porque

Los solteros son hombres

es analíticamente verdadera mientras que

Los solteros son histéricos

es sintéticamente verdadera: si son dos factores independientes, ¿por qué una característica que es propia de uno habría de determinar alguna diferencia en el otro? (Del mismo modo, desde esta perspectiva,

90. Véase Quine 1951. En síntesis, Quine ataca por dos flancos a la distinción analítico/sintético: por un lado, la considera una noción semántica oscura e inexplicable; por otro, la acusa de formar parte de una visión fracasada del conocimiento y la ciencia (el fenomenalismo o verificacionismo reduccionista).

91. Este parágrafo está inspirado en un comentario de Devitt. Véase al respecto Devitt 1996: cap. 1.

afirmar que "histérico" forma parte del significado de "soltero" no implica afirmar nada sobre el valor de verdad de "Los solteros son histéricos").

Sea por las razones que fueren, los teóricos del rol conceptual no suelen comprometerse con la distinción analítico-sintético.[92] No hay duda entonces de que el tipo de teoría examinada involucra un compromiso con el holismo semántico.[93] [94] Esto no implica que sus defensores hagan explícito este compromiso; por el contrario, muchos de ellos no sólo no lo hacen explícito sino que incluso lo niegan.[95] Precisamente, uno de los problemas de la tesis holista es que muy pocas veces aparece claramente formulada.[96] Ahora bien, ¿qué consecuencias tiene un compromiso de este tipo? En lo que sigue, me concentraré entonces en la evaluación de esta cuestión. Según sus defensores, el holismo presenta la ventaja de otorgarle a la teoría semántica un máximo poder explicativo de la conducta humana, puesto que la explicación de la conducta requiere la adscripción de

92. Véase, por ejemplo, Block 1986: p. 629.
93. Cabe aclarar que la tesis conversa no es válida: es posible ser holista semántico sin suscribir la semántica de roles conceptuales ni la teoría del doble factor, como es el caso de Davidson, defensor de una versión peculiar o heterodoxa de la teoría correspondentista de la verdad, y el de Quine, un verificacionista.
94. Por otro lado, si uno suscribe algún tipo de teoría causal, según la cual el significado se identifica con la referencia, esto es, la relación causal directa con un objeto o propiedad del mundo, creerá probablemente que ninguna propiedad inferencial/creencia asociada constituye el significado de una expresión, es decir, será un atomista semántico. Por último, si uno suscribe una teoría según la cual el significado de la mayor parte de las expresiones de un lenguaje es la referencia, pero ésta no sólo está constituida por la relación causal sino también por ciertas descripciones que contribuyen a su fijación, se verá comprometido con la creencia de que sólo algunas propiedades inferenciales/creencias asociadas constituyen el significado, es decir, con una teoría semántica molecularista.
95. Véase al respecto Fodor & Lepore 1992: cap. 6.
96. Este problema ha sido señalado tanto por Fodor & Lepore (Fodor & Lepore 1992) como por Devitt (Devitt 1996). Devitt se ocupa, sin embargo, de reproducir varias citas en donde la tesis aparece formulada de una manera clara y contundente. Véase Devitt 1996: pp. 15-16.

significados finamente discriminados y los significados holísticos, altamente individualistas, poseen el suficiente grado de discriminación. Retomando el ejemplo anterior, la negativa de Pedro a ir al concierto puede ser explicada por su creencia de que tocará no sólo un cierto individuo en particular sino un individuo que él se representa como el primer violinista del Cuarteto local o, en otros términos, que él asocia con un concepto, EL PRIMER VIOLINISTA DEL CUARTETO LOCAL, que desempeña cierto rol o cierta función en su sistema conceptual (si se lo representara de distinto modo o lo asociara con otro concepto como, por ejemplo, EL RECIENTE GANADOR DEL PREMIO BACH, su actitud habría sido distinta). En general, a quién se refiere el agente no importa tanto para la iluminación de sus motivos para actuar como de qué manera se representa a la persona referida: a los fines de la explicación psicológica, parece entonces necesario concentrarse en el rol funcional que los diferentes conceptos, asociados con distintas expresiones, juegan en la psicología del hablante/agente. Volviendo al ejemplo, los conceptos EL PRIMER VIOLINISTA DEL CUARTETO LOCAL y EL RECIENTE GANADOR DEL PREMIO BACH juegan roles distintos en la psicología de Pedro y son, por tanto, semánticamente distintos, a pesar de hacer referencia al mismo individuo. En otras palabras, la diferencia en cuestión es perfectamente captada por un significado individualizado en función de propiedades inferenciales, tal como el rol conceptual, pero no lo es por un significado individualizado en función de la relación causal con un individuo, tal como el referente.

Sin embargo, cabe señalar lo siguiente. En primer lugar, para lograr el objetivo señalado, a saber, la explicación de la acción o la conducta, si bien se requiere indudablemente adscribir significados más finamente discriminados que los referentes, no parece ser necesario adscribir significados tan finamente discriminados como los roles conceptuales. Volviendo al ejemplo, es cierto que atribuir a Pedro una creencia que involucre una relación referencial directa con un individuo específico no es suficiente para lograr la explicación deseada, pero ¿implica eso que es necesario atribuirle una creencia con un cierto rol inferencial, con la naturaleza holística que le es

propia? Sin duda, si el objeto es distinguir entre EL PRIMER VIOLINISTA DEL CUARTETO LOCAL y EL RECIENTE GANADOR DEL PREMIO BACH (cuando, por hipótesis, ambos conceptos refieren al mismo individuo), la consideración tanto de la totalidad de las respectivas propiedades inferenciales como de la totalidad de las creencias asociadas parece por lo menos excesiva: parece mucho más razonable hacer la distinción en cuestión mediante algo similar al recurso fregueano de apelar al modo de presentación del referente. En otros términos, la mayor parte de las oraciones que mantienen relaciones inferenciales con las oraciones que contienen los conceptos en cuestión e incluso gran parte de las oraciones que contienen los conceptos en cuestión (tales como "Los pinos no pierden sus hojas" o "El tío del violinista del cuarteto local es vecino de mi madre" respectivamente) no parecen desempeñar rol alguno en la explicación de la conducta de Pedro; en términos más simples, la mayor parte de las creencias de Pedro son absolutamente irrelevantes para explicar su decisión de no ir al concierto. Sin embargo, comprometerse con roles conceptuales implica, en virtud del carácter holístico de los mismos, involucrar a todas esas relaciones inferenciales y a todas esas creencias, claramente irrelevantes, en la explicación de porqué Pedro ha decidido no ir al concierto esta noche.

En segundo lugar, es preciso señalar un problema aun más grave: la atribución de roles conceptuales no sólo no es necesaria sino que ni siquiera es posible. En general, se considera que la consecuencia negativa más importante que acarrea el holismo es la pérdida de estabilidad del significado. El holismo convierte al significado en una entidad completamente individual: no es posible que dos personas compartan un mismo significado (a menos que sean duplicados funcionales) porque ningún término tiene las mismas propiedades inferenciales para dos personas distintas. A modo de ejemplo, desde el punto de vista de las teorías comprometidas con roles conceptuales, la creencia de María, habitante del hemisferio sur, de que la Cruz del Sur es la constelación más brillante forma parte del significado de CRUZ DEL SUR tal como se instancia en su mente, si bien no en la de Mathew, habitante del hemisferio norte; su tendencia a recostarse en el pasto para mirar el cielo, asociada con el

concepto en cuestión, también forma parte del significado de su concepto, pero no del de Mathew, que vive rodeado de nieve. El holismo implica entonces la renuncia al intento de identificar el rol funcional de un concepto con algo compartido por todos los miembros de una cierta comunidad lingüística. Resulta imposible que dos personas tengan representaciones con el mismo rol conceptual; más aun, es incluso imposible que una representación mantenga un cierto rol conceptual a lo largo del tiempo, puesto que siempre habrá algún cambio en la vida mental del individuo. En términos del ejemplo anterior, dado que María adquirirá nuevas creencias y nuevos patrones de conducta, todos ellos deberán ser incluidos en sus futuros conceptos. De acuerdo con esto, el significado deja de ser tanto intrasubjetivo como público. En síntesis, los roles conceptuales, en la medida en que son altamente individualistas, involucran un serio problema de individualización o identificación: no hay un criterio claro para distinguir un rol conceptual de otro.[97]

La falta de estabilidad torna igualmente imposible la apelación al principio de composicionalidad para explicar ciertos rasgos de la comprensión lingüística, tales como la productividad y la sistematicidad.[98] De acuerdo con el principio en cuestión, el significado de toda oración del lenguaje depende exclusivamente de los significados de las palabras que la componen y de su estructura sintáctica. Pero los roles conceptuales no son composicionales: el rol conceptual de una oración no es producto de los roles conceptuales de sus partes constitutivas. A modo de ejemplo, el rol conceptual de

97. Véase Fodor & Lepore 1992: p. 169. Véase también Block 1986: p. 629.
98. De acuerdo con la primera, es posible comprender un número infinito de oraciones del lenguaje natural a partir del dominio de un vocabulario finito y un número finito de reglas sintácticas. La sistematicidad, por su parte, es el rasgo según el cual, si es posible comprender la oración *p*, entonces también será posible comprender muchas oraciones conceptualmente cercanas a *p*, lo cual puede ilustrarse a través del siguiente ejemplo: la comprensión de una oración como "Juan come una manzana"implica la comprensión de otras oraciones tales como "una manzana es comida por Juan", "algo es comido por Juan", etc., lo cual pone de manifiesto que la comprensión no se refiere a fragmentos aislados del lenguaje.

El mar es azul

no sólo depende de las creencias asociadas con "mar" y "azul" respectivamente sino también de las creencias asociadas con "mar azul" y, en última instancia, de todas las creencias expresadas en el lenguaje. Esto parece ser especialmente grave, dado que no se ha ofrecido ninguna explicación alternativa de los fenómenos mencionados, esto es, la productividad y sistematicidad lingüísticas.

Es interesante destacar que los teóricos del doble factor son conscientes de la falta de estabilidad de los roles conceptuales pero no consideran que esto constituya un problema grave. En su opinión, no importa cuán subjetivos e inestables sean los roles conceptuales, el carácter público e intersubjetivo del significado es preservado por el otro factor, el factor referencial.[99] En este punto, sólo quiero, sin embargo, consignar lo siguiente. Dados dos factores tan distintos, uno de los cuales es de carácter holístico y el otro no, ¿cúal es el sentido de considerarlos como dos constituyentes independientes del significado? ¿Qué los hace dos aspectos de la misma cosa en lugar de dos cosas distintas? En otras palabras, ¿cuál es la razón para considerar que los roles conceptuales son parte del significado de los términos del lenguaje? Como se vio anteriormente, si bien el factor referencial parece ser claramente necesario para el logro de ciertos propósitos semánticos, no parece haber ningún propósito semántico que requiera la postulación de roles conceptuales -con la naturaleza holística que les es propia. Estos últimos, al igual que las ideas que Frege insiste en diferenciar de los sentidos, incluyen demasiadas relaciones, muchas más de las comúnmente involucradas en las usuales adscripciones de significado.[100] Intuitivamente, no parece haber razón alguna para incluir tales relaciones en la noción de significado -lo que parece sugerir

99. Véase, por ejemplo, Field 1977: p. 398.
100. En palabras de Frege: "Un pintor, un jinete y un zoólogo probablemente asocien distintas ideas con la palabra "Bucéfalo". Esto constituye una diferencia esencial entre la idea y el sentido de un signo, el cual puede ser propiedad común de muchas personas, y no es, por tanto, una parte o característica de la mente individual." Frege 1892: p. 160.

claramente que el carácter público e intersubjetivo es una nota esencial de la noción preteórica de significado.

En síntesis, por un lado, no hay evidencia de que los sentidos y los referentes sean conceptualmente distintos y mutuamente independientes; pero los roles conceptuales son, por definición, conceptualmente distintos e independientes de los referentes. Por otro, no hay evidencia de que modos holísticos jueguen rol alguno en semántica; pero los roles conceptuales son holísticos. Ambas tesis parecen indicar que los sentidos no son roles conceptuales. Finalmente, también es preciso volver a señalar que estas teorías no ofrecen ninguna solución al otro problema que, según vimos, amenaza a las teorías causales de la referencia directa, a saber, el problema del respecto. En este sentido, no representan ningún avance significativo por sobre la teoría de Kripke. Por estas razones, la explicación del significado propuesta por las teorías del doble factor no resulta, en mi opinión, en absoluto convincente.

2. La teoría descriptivo-causal de Devitt y Sterelny

2.1 Las tesis principales

2.1.1. Los sentidos como tipos de cadenas designativas

La discusión anterior parece dejar en claro un punto: para dar cuenta del significado de manera adecuada, es necesario asignar no sólo la propiedad semántica de referir a un objeto (la referencia) sino alguna otra propiedad semántica, más fina que la anterior, lo que Frege denominó "sentido" o "modo de darse del objeto". Otro punto quedó claro también o, por lo menos, he intentado que así sea: no es adecuado concebir a los modos en cuestión como estereotipos o roles conceptuales. ¿Significa esto acaso que, pese a los problemas antes mencionados, es preciso retornar al modelo descripcional, en alguna de sus versiones? ¿No eran acaso los problemas de la ignorancia y el error y la esencial incompletitud de las teorías

descripcionales, aun en sus variantes más refinadas, razones contundentes para abandonarlas y adoptar en su reemplazo el nuevo modelo causal? Si uno de los problemas principales de éste último está determinado por la necesidad de incorporar una nueva propiedad semántica que permita dar cuenta de los problemas semánticos tradicionales y otorgue al mismo tiempo a la teoría poder explicativo de la conducta, ¿hay acaso algún modo de hacerlo sin abandonar el modelo causal? En este punto, resulta oportuno introducir una nueva propuesta, a saber, la teoría descriptivo-causal de Devitt y Sterelny.

Esta teoría ejemplifica la concepción causal pero no es una teoría de la referencia directa, como la de Kripke, ni incluye como parte una teoría de la referencia directa, como la de Putnam. De este modo, se incorpora, junto a la propiedad semántica de referir a un objeto, la de hacerlo de determinada manera, pero ésta última es concebida no en términos de una descripción asociada sino en términos del tipo de cadena causal que conduce al referente -lo que los autores denominan "tipo de cadena designativa" ("*type of designational chain*").[101] A modo de ejemplo, desde este punto de vista, tanto "Charles Dodgson" como "Lewis Carroll" están causalmente fundados en el mismo individuo, por lo que ambos nombres tienen el mismo referente; sin embargo, en la medida en que el tipo de cadena causal que conduce al referente es en cada uno

101. Cabe señalar que estos autores dan un sentido más preciso a conceptos que hasta aquí se han usado de manera intercambiable, tales como los conceptos de referencia, denotación y designación. Para Devitt y Sterelny, el concepto de referencia es general y abarca a los otros dos, que serían sus dos variantes principales: un término *denota* a un objeto si y sólo si refiere a él en virtud de connotar ciertos predicados que se aplican al objeto de manera única; por otro lado, un término *designa* un objeto si y sólo si refiere a él en virtud de una cadena causal que conduce al objeto en cuestión. En síntesis, la referencia puede darse en virtud de la aplicación única de predicados (denotación) o en virtud de la existencia de una cadena causal apropiada (designación).

de estos casos distinto, es posible considerar que los nombres en cuestión tienen distintos sentidos.[102]

Ahora bien, cabe aclarar que las cadenas designativas que subyacen a los usos de distintos términos forman distintas redes que son mentalmente procesadas por la comunidad lingüística.[103] En otros términos, parte de las cadenas designativas que constituyen las redes correspondientes a los distintos términos se encuentran en la mente de los individuos, forman parte de sus creencias y procedimientos de categorización. De este modo, el sentido de un término puede ser identificado con la red causal correspondiente, que vincula en definitiva al referente del término en cuestión con cierto procesamiento que tiene lugar en la mente del sujeto. En términos de Devitt:

> *Las cadenas-d (d-chains) son del tipo que nos interesa sólo si están causalmente vinculadas por el procesamiento mental de la comunidad lingüística en una red única de fundaciones y transmisiones de la referencia.*[104]

Sin intentar profundizar en los detalles, que conducen a múltiples refinamientos, lo que me interesa destacar es la idea general de

102. La definición de sentido ofrecida por Devitt y Sterelny pone de manifiesto la distinción *tipo-caso*: a todos los casos o ejemplares del mismo tipo de nombre, tal como "Charles Dodgson", les corresponde el mismo tipo de sentido, esto es, el mismo tipo de cadena designativa. Esto hace patente, a su vez, lo que se presupuso a lo largo de todo este trabajo, a saber, que cuando se habla de las expresiones y de sus significados, uno se refiere directamente a tipos -y sólo indirectamente a los casos de estos tipos. Cabe aclarar que no hay aquí implicado un compromiso ontológico con entidades abstractas, puesto que el discurso acerca de tipos puede interpretarse como un modo más simple de hablar acerca de varios individuos particulares. Para una explicación muy clara de la distinción tipo-caso, véase, por ejemplo, García-Carpintero 1996: cap. 1.
103. De ahí que las condiciones de identidad para la red correspondiente a una cierta expresión-*tipo* se den en términos de las cadenas designativas que involucran *casos* del tipo en cuestión unidos entre sí en virtud del procesamiento mental llevado a cabo por la comunidad de hablantes.
104. Devitt 1996: p. 167.

que el pensamiento y el lenguaje se fundamentan en múltiples redes de cadenas causales. No todos los miembros de la comunidad lingüística comparten las mismas redes puesto que no todos ellos están en contacto causal con los mismos tipos de cadenas designativas. En términos del ejemplo anterior, bien puede ocurrir que alguien esté en contacto causal con uno de los tipos de cadenas que conduce a Lewis Carroll (como, por ejemplo, la que subyace a "Lewis Carroll") pero no con otros (como, por ejemplo, la que subyace a "Charles Dodgson").

Por consiguiente, por un lado, los tipos de cadenas designativas no son modos descriptivos sino modos causales; de este modo, esta teoría difiere radicalmente de la teoría descripcional en lo que respecta a la concepción del sentido. Si bien Devitt y Sterelny, como veremos más adelante, dan un rol a las descripciones asociadas en la determinación de la referencia de un término, no identifican al conjunto de dichas descripciones con el sentido del término. El sentido de un término no es una representación conceptual (como en Frege) ni un conjunto de representaciones conceptuales (como en Searle), no es una entidad psicológica o mental de ningún tipo, no es un objeto de conocimiento: es un tipo de cadena causal. Lo que se concibe como algo (en su mayor parte) externo al sujeto no es sólo la dimensión referencial o el referente de un término sino su sentido mismo.

Por otro lado, como se señaló previamente, al igual que los sentidos tanto clásicos como refinados y a diferencia de los roles conceptuales, los tipos de cadenas designativas determinan la referencia. Se trata, por tanto, de una teoría representacionalista del significado: el rol representacional del significado es el rol referencial -lo que no quita que éste esté en parte determinado por propiedades inferenciales. La concepción representacionalista del significado respeta el *dictum* fregueano según el cual el sentido de una expresión es lo que determina su referencia.

Como puede preverse, la teoría presentada permite ofrecer una solución razonable a los problemas tradicionales. Piénsese, por ejemplo, en el problema de la identidad: en el marco de esta teoría, es perfectamente posible explicar porqué

Lewis Carroll es Charles Dodgson

es una oración informativa, mientras que

Lewis Carroll es Lewis Carroll

no lo es. Como señalé anteriormente, "Lewis Carroll" y "Charles Dodgson" tienen distintos significados porque, si bien ambos nombres designan al mismo individuo, el tipo de cadena designativa que subyace al primer nombre es distinto del que subyace al segundo. En otros términos, la emisión de la primera oración involucra el acceso a dos tipos distintos de cadenas causales y, por tanto, el ejercicio de dos habilidades semánticas distintas, mientras que la emisión de la segunda involucra el acceso a un único tipo de cadena causal y el ejercicio de una sola habilidad semántica -junto con la creencia en el principio de identidad.

Es asimismo posible ofrecer una explicación del hecho de que

Pedro cree que Lewis Carroll es un escritor inglés del siglo XIX

sea verdadera mientras

Pedro cree que Charles Dodgson es un escritor inglés del siglo XIX

sea falsa: la razón de ello es que, aun cuando "Lewis Carroll" y "Charles Dodgson" designen al mismo individuo, los tipos de cadenas designativas subyacentes difieren y Pedro tiene acceso a uno de ellos (el tipo subyacente a "Lewis Carroll") pero no al otro (el tipo subyacente a "Charles Dodgson").

En cuanto al problema de las oraciones que contienen nombres ficticios o vacíos, la teoría sostiene que lo que ocurre es que sus respectivos tipos de cadenas designativas no están fundados apropiadamente en objetos, ya sea porque el hablante cree percibir algo (o tal vez incluso perciba algo) cuando de hecho no hay nada que sea

percibido (como es el caso de las alucinaciones), ya sea porque los objetos en cuestión no existen o no son objetos espacio-temporales (como es el caso de los personajes de las obras de ficción).

Del mismo modo, el problema de las oraciones existenciales negativas encuentra una solución: quien afirma, por ejemplo,

Palas Athenea no existe

está afirmando una oración significativa y verdadera; puesto que, por un lado, el nombre utilizado tiene un tipo de cadena causal subyacente (esto es, un sentido) y, por el otro, la cadena causal en cuestión no está fundada en ningún individuo realmente existente (es decir, ningún referente).

2.1.2. El rol de las descripciones

Cabe destacar que la presente teoría también intenta ofrecer una solución al otro problema mencionado, a saber, el problema de ambigüedad denominado "problema del respecto" ("*qua problem*"). La solución en cuestión apela a la reintroducción de descripciones en el bautismo inicial, de modo tal de seleccionar uno entre los múltiples aspectos del objeto como correlato ontológico de la representación involucrada. Ahora bien, las descripciones postuladas no son del tipo de las que proporcionan conocimiento identificador del referente o, en otros términos, un conjunto de condiciones necesarias y suficientes para la aplicación del término a un objeto.

Por un lado, se considera que cuando alguien bautiza a un objeto individual con un nombre propio, ha de tener cierta actitud o ciertas creencias respecto de lo que está bautizando, por ejemplo, ha de considerarlo un individuo especial y no un ejemplar cualquiera de una clase natural, un animal y no una parte no separada de un animal. A modo de ejemplo, si se quiere fundar el nombre "Oliverio" en un determinado gato, no basta con tener un contacto perceptivo directo con él -a partir del cual se constituye la relación causal- sino que es también necesario disponer de cierto conocimiento general que haga posible describir al gato en cuestión mediante ciertos términos

categoriales, tales como "individuo" y "animal": sólo de esa manera "Oliverio" quedará causalmente fundado en el gato *qua* individuo y no *qua* miembro o ejemplar de una clase natural, a saber, la especie Gato; del mismo modo, se asegura que el nombre en cuestión quede fundado en el gato *qua* objeto completo y no en una parte no separada o en un estadio temporal de un gato. En términos de los autores:

> *La consideración de 1 [el problema del qua] muestra que debe haber algo acerca del estado mental del fundador que determine que el nombre está fundado en la causa de la experiencia perceptiva qua objeto completo. No ayuda en nada decir que es la intención del fundador lo que hace que ése sea el caso. ¿En virtud de qué el fundador tiene intención de referirse al objeto completo? Parecería que el fundador debe, en cierto nivel, pensar la causa de su experiencia bajo cierto término categorial general tal como "animal" u "objeto material". Es por esta razón que la fundación es en Nana [una cierta perra] y no en una parte espacio-temporal de aquélla.*[105]

Por otro lado, si se trata del bautismo de una clase natural, como, por ejemplo, la especie Gato, se considera que el hablante debe pensar al objeto percibido no sólo como ejemplar de una clase natural (y no como individuo o ejemplar de una clase no natural) sino también como ejemplar de una determinada clase natural, pues todo ejemplar lo es de muchas clases naturales distintas; por consiguiente, también en este tipo de casos debe ser capaz de utilizar ciertos términos para describir la clase en cuestión, que permitan distinguir su naturaleza de la de toda otra clase. En términos de Devitt y Sterelny:

> *En suma, el fundador de un término de clase natural asocia, consciente o inconscientemente, con ese término, primero, alguna descripción que en efecto clasifica al término como un término de clase natural; segundo, algunas descripciones que determinan qué naturaleza de la muestra es pertinente para la referencia del término.*[106]

105. Devitt & Sterelny 1987: pp. 64-65.
106. Devitt & Sterelny 1987: p. 74.

En síntesis, ya se trate de un nombre propio o de un término general, en el bautismo inicial, el hablante debe poseer cierta capacidad descriptiva respecto del objeto involucrado.

La presente propuesta involucra por tanto una teoría mixta o descriptivo-causal de la fijación de la referencia: en el mecanismo de fijación, interviene tanto la relación causal como el uso de (ciertas) descripciones. En otras palabras, la relación referencial resulta estar determinada no sólo por la relación causal con un objeto sino también por ciertas descripciones del objeto en cuestión, las cuales sirven para caracterizar el estado psicológico de quienes fundan la representación en el objeto. En otras palabras, hay una relación causal constitutiva, pero ésta sólo puede ser identificada mediante la apelación a las creencias del sujeto respecto de aquello de lo que se habla, lo cual involucra de manera esencial el uso de descripciones.

Ahora bien, Devitt y Sterelny consideran que la explicación descriptivo-causal de cómo se determina la referencia en el bautismo inicial debe ser acompañada por una explicación puramente causal de cómo la referencia es transmitida a los otros miembros de la comunidad lingüística. En otros términos, todo miembro de la comunidad ausente en el bautismo, en la medida en que participe de la red causal correspondiente, puede usar la representación exitosamente, aun cuando sea completamente ignorante o esté completamente equivocado acerca de todas las propiedades -incluidas las generales o categoriales- del objeto referido o representado. En los usos posteriores, es posible tomar prestado el término de quienes estaban presentes en el bautismo, lo cual sólo requiere una apropiada inserción en la red causal correspondiente. De este modo, se considera que las descripciones, que, como hemos visto, cumplen un rol vital en la fijación de la referencia, no juegan rol alguno en el mecanismo de transmisión de la misma.

2.2. El análisis crítico

Considero que, en general, una teoría descriptivo-causal como la propuesta por Devitt y Sterelny puede funcionar tanto para los nombres propios como para los términos generales. Sin embargo,

hay ciertos aspectos de la teoría defendida por estos autores que pueden, en mi opinión, ser cuestionados. Mi crítica se articula fundamentalmente en dos partes, que están estrechamente relacionadas entre sí: la primera está centrada en la solución propuesta al problema del sentido; la segunda, se ocupa del intento de los autores por solucionar el problema de la ambigüedad. El desarrollo de cada una de estas partes es, entonces, a la vez causa y ocasión de la exposición de la variante descriptivo-causal que intento defender.

2.2.1. La respuesta al problema del sentido

Estoy de acuerdo con Devitt y Sterelny en que los sentidos pueden ser definidos en términos de los tipos de cadenas causales que conducen a los referentes y que pueden, por consiguiente, ser considerados en gran parte externos a los hablantes. Sin embargo, no coincido con la tesis de que los sentidos son puramente causales o no son en absoluto descriptivos, es decir, la tesis de que son completamente externos a los hablantes. La argumentación a favor de esta postura consta de dos partes principales: en una primera parte, critico la tesis, compartida por Kripke, de que las descripciones utilizadas en la fijación de la referencia no son constitutivas del significado -crítica que fue inicialmente formulada en relación con la teoría de Kripke; en la segunda parte, critico la tesis, también presente en Kripke, de que el mecanismo de transmisión de la referencia es puramente causal. El abandono de ambas tesis permite, en mi opinión, sostener una concepción descriptivo-causal de los sentidos, que resulta no sólo más rica sino más acorde con la evidencia disponible.

En primer lugar, cabe preguntarse entonces porqué Devitt y Sterelny consideran, como hace Kripke, que las descripciones que contribuyen a fijar la referencia de un nombre o de un término general no forman parte de su significado; esto es, por qué razón separan, a la manera de Kripke, la teoría de la fijación de la referencia de la teoría del significado. Como se recordará, esta estrategia fue criticada a propósito del análisis de la teoría de Kripke; el argumento en contra de ella no es por tanto distinto del utilizado en esa instancia. En síntesis, si se considera que el mecanismo de fijación de la

referencia es independiente del significado, no puede esperarse que la explicación de la referencia arroje luz alguna sobre la naturaleza del significado. Pero, entonces, ¿cuál podría ser el propósito de ocuparse de la referencia? Desde mi punto de vista, el interés por la referencia sólo puede basarse en el interés por explicar el significado. En otras palabras, si se parte de la base, común a todas las teorías referencialistas del significado, de que el significado de un término está constituido, por lo menos en parte, por aquello que determina su referencia, no parece tener sentido alguno excluir del significado a los elementos claves de los mecanismos referenciales. Si esto es así, parece más sensato considerar que las descripciones, consideradas un elemento clave para la fijación de la referencia, forman parte del significado de los términos.

En este punto, es preciso hacer la siguiente aclaración. De ninguna manera intento defender la tesis de que el hecho de que las descripciones intervengan en la fijación de la referencia de los nombres propios y términos generales fundamenta el hecho de que unas y otros tienen el mismo significado o son semánticamente equivalentes -tesis contra la cual Kripke argumenta extensamente. La tesis defendida es, por el contrario, que el significado de los nombres propios y términos generales está *parcialmente* constituido por el conjunto de descripciones que contribuyen a la fijación de su referencia.

Más aún, si no se acepta la tesis de que las descripciones en cuestión son constitutivas del significado, no es posible considerar que la apelación al uso de descripciones en el bautismo inicial permite ofrecer una solución al problema de la ambigüedad -como sin duda hacen Devitt y Sterelny. La razón de ello es que, en mi opinión, éste último no es un problema que afecta a la referencia pero no al significado sino, por el contrario, es un problema que afecta a la referencia en tanto constitutiva del significado. Dicho brevemente, el problema es que el significado, definido en términos de referencia, definida a su vez en términos de causalidad, es ambiguo; eliminar la ambigüedad implica incorporar en la definición algo más, ajeno a la noción de causalidad, como son las descripciones. De este modo, no incorporar las requeridas descripciones en la definición del significado implica no solucionar el problema de ambigüedad. Ejemplo de ello

es, en mi opinión, la teoría del propio Kripke, a la cual, como vimos, es posible plantear el problema del respecto -problema, que como también vimos, no es realmente tenido en cuenta ni evaluado por el autor. A diferencia de Kripke, Devitt y Sterelny parecen tomar en cuenta el problema en toda su extensión; sin embargo, la similaridad entre su posición y la posición kripkeana torna imposible considerar que la primera contiene realmente una respuesta al problema del respecto. Como se recordará, Kripke afirma de múltiples maneras, en repetidas ocasiones, que las descripciones pueden ser usadas para fijar la referencia de nombres y términos generales pero no dan su significado; Devitt y Sterelny restringen de la misma manera el rol de las descripciones; luego, ¿por qué habría de considerarse que los últimos, pero no el primero, aportan con tal tesis alguna solución al problema de la ambigüedad?

En síntesis, desde mi punto de vista, en primer lugar, hay razones suficientes para considerar que las descripciones utilizadas en el bautismo inicial no sólo determinan la referencia del término involucrado sino que también constituyen parcialmente su significado, más específicamente, su sentido. De ahí, en parte, que no crea que los sentidos puedan ser definidos *exclusivamente* en términos de cadenas causales.

Paso ahora a la consideración del segundo punto antes mencionado, a saber, la defensa de un mecanismo puramente causal de transmisión de la referencia. Como se ha visto, Devitt y Sterelny consideran que la explicación descriptivo-causal de cómo se determina la referencia en el bautismo inicial (es decir, del mecanismo de fijación de la referencia) debe ser acompañada por una explicación puramente causal de cómo la referencia es transmitida a los otros miembros de la comunidad lingüística (es decir, del mecanismo de transmisión de la referencia). En otros términos, todo miembro de la comunidad ausente en el bautismo, en la medida en que participe de la red causal correspondiente, puede usar la representación exitosamente, aun cuando sea completamente ignorante o esté completamente equivocado acerca de todas las propiedades -incluidas las generales o categoriales- del objeto referido o representado. Ahora bien, dudo de que se trate de una tesis psicológica y psicolingüísticamente plausible, lo cual, en mi opinión, resultaría apro-

piado, puesto que se trata de una afirmación acerca de lo que los hablantes saben sobre el individuo referido.

En mi opinión, los estados epistémicos que pueden requerirse de los hablantes/pensantes competentes presentan distintos grados. Pienso que es imposible exigirles la posesión de conocimiento de condiciones necesarias y suficientes para la aplicación de los términos, esto es, un estado epistémico del máximo grado, como sugieren las tradicionales teorías descripcionales de la transmisión de la referencia. Sin embargo, puede ser plausible requerir de aquéllos un estado epistémico menos estricto, tal como el conocimiento de que el objeto referido satisface, por cierto no de manera única, alguna descripción definida o indefinida. Tal vez pueda pensarse incluso que el estado epistémico en cuestión no constituye conocimiento o, por lo menos, que no se trata de conocimiento consciente: el sujeto no sabe que sabe ni siquiera cree que sabe sino que tan sólo cree algo acerca del objeto del cual habla -algo que sin duda es verdadero del mismo pero el sujeto no está en posición de decir que lo es o, en otras palabras, no tiene justificación alguna para hacer una afirmación semejante. Lo que dicen Devitt y Sterelny con respecto al estado epistémico involucrado en el bautismo inicial, se debería extender, desde mi punto de vista, al estado epistémico que caracteriza a los usos posteriores:

> *Estamos comprometidos con la asociación de descripciones determinantes de la referencia con los términos por parte de los hablantes (...) En las teorías tradicionales, esta suerte de asociación significaba que el hablante tenía conocimiento. ¿Deberíamos acaso continuar con esta tradición? Hay una consideración de peso en contra de ello. Como veremos más tarde (...) hay razones sugestivas si bien no decisivas para ubicar nuestra competencia lingüística en un subsistema psicológico, un módulo. Una característica de ese módulo es que la información contenida en él no es accesible de manera automática y completa al procesador central del hablante. Si las asociaciones están modularizadas, entonces no constituyen conocimiento. Son información en el hablante, no del hablante.*[107]

107. Devitt & Sterelny 1987: p. 79.

De acuerdo con esto, a modo de ejemplo, a fin de tomar prestado el nombre "Oliverio" para un cierto gato, se requiere no sólo tener un adecuado contacto causal con quienes estuvieron presentes en su bautismo sino también poseer alguna descripción tal como "el gato siamés de mi amiga", que puede no seleccionar a un objeto único: en otras palabras, a pesar de que mi amiga tenga varios gatos siameses, es plausible pensar que el uso de la descripción anterior junto con la relación causal correspondiente permiten seleccionar de manera unívoca a uno de ellos como el referente de "Oliverio". Para poner otro ejemplo, a fin de tomar prestada la palabra "eau" de mis amigos franceses, necesito no sólo mantener una apropiada relación causal con ellos sino también creer conscientemente, por ejemplo, que el eau es un cierto tipo de sustancia natural que calma la sed. La ignorancia total, que implica una ausencia total de creencias sobre aquello de que se habla, parece altamente implausible, mientras que el error total sólo parece plausible si es masivamente compartido por la comunidad. Debe tomarse en cuenta, sin embargo, que esto dista mucho del requisito descripcional clásico de conocimiento identificatorio del referente en términos de condiciones necesarias y suficientes para la aplicación de una representación, y está más cerca en cambio de la idea principal que subyace a la teoría del cúmulo de descripciones.

Tal vez pueda pensarse que existe una historia de usos individuales de ciertas representaciones que es paralela a su historia social. En términos más específicos, el momento en el cual un cierto individuo adquiere una representación recrea, hasta un cierto punto, el bautismo inicial, es decir, el momento en el cual la sociedad en su conjunto la adquiere. En ambos momentos, los hablantes usualmente tienen un contacto perceptivo directo con el referente, el cual es posible en parte en virtud de su dominio de ciertas descripciones, previamente adquiridas. Los usos ulteriores de la representación fuera de la situación de aprendizaje son similares a los usos de quienes toman prestada la referencia: el individuo no necesita entonces tener contacto perceptivo directo con el objeto referido; pero, además de mantener una apropiada relación causal con él, necesita poseer algunas descripciones.

En síntesis, en segundo lugar, considero que las descripciones de todo tipo (esto es, no sólo las categoriales) son tan importantes en la historia del uso de una palabra como la cadena causal misma. La diferencia entre ambos factores reside en la mayor estabilidad de la cadena causal respecto de las descripciones involucradas. Sin embargo, las descripciones son necesarias para seleccionar la cadena causal a la que pertenece el uso de determinada palabra. Como consecuencia de ello, no considero que el mecanismo de transmisión de la referencia sea, para los nombres propios y términos generales, muy distinto del mecanismo de fijación: en ambos casos, se trata de una relación descriptivo-causal con un objeto y la relación en cuestión es parcialmente constitutiva del sentido de las expresiones involucradas.

De este modo, pienso que las descripciones que he propuesto asociar con toda representación (no básica) tanto en el bautismo inicial como en los usos posteriores pueden ser consideradas como parte de su sentido. Mi razón en apoyo de ello es el hecho de que esas descripciones determinan la manera en que la representación en cuestión es procesada y almacenada en nuestra mente; por consiguiente, parecen formar parte de lo que Devitt llama "red causal" o "tipo de cadena designativa". Por ejemplo, si con "Venus" asocio la descripción "un planeta", será almacenada de manera distinta -en relación, por ejemplo, con creencias astronómicas- que si con ELLA asocio "una diosa romana" o "el perro de mi vecino" -en relación, en cambio, con creencias mitológicas y deseos de matar, respectivamente-. En otras palabras, las redes causales subyacentes a nuestros usos lingüísticos, en la medida en que incluyen un procesamiento subjetivo, parecen involucrar cierta información (que no puede caracterizarse como conocimiento de características identificatorias aplicables a determinados objetos de manera exclusiva) y tal vez ni siquiera como conocimiento (por lo menos, no como conocimiento consciente) acerca de los objetos referidos. De acuerdo con esto, no creo que los sentidos o modos de presentación puedan ser considerados puramente causales -lo cual puede hacerse si se piensa, como Devitt y Sterelny, por un lado, que las descripciones que contribuyen a fijar la referencia de un determinado término son ajenas a su sentido, y, por otro, que el mecanismo de transmisión de la referencia es a su

vez puramente causal. Lo que intento defender es entonces, como mencioné al comienzo de este apartado, la necesidad de postular una noción de sentido más rica que la propuesta por Devitt y Sterelny.

Tal vez pueda pensarse que mi caracterización de los sentidos se acerca demasiado a la tradición fregueana. Sin embargo, hay dos puntos que quisiera enfatizar. Por un lado, las descripciones aquí involucradas están acompañadas por una cadena histórico-causal, no están solas en el cumplimiento del rol de fijar y transmitir la referencia. Por otro lado, y quizás como consecuencia del punto anterior, no nos proveen de conocimiento que nos permite identificar al referente de manera unívoca sino de algunas creencias de distinto grado de especificidad, tal vez inconscientes pero, por lo general, socialmente compartidas y perceptualmente adecuadas, acerca de aquél.

2.2.2. La respuesta al problema del respecto

Ante todo, cabe reiterar lo expresado más arriba: en mi opinión, el rol asignado por la teoría de Devitt y Sterelny a las descripciones, a saber, fijar la referencia sin constituir el significado de los respectivos términos asociados, impide, por las razones consignadas, considerar que se haya ofrecido una solución al problema de la ambigüedad. La solución en cuestión requiere que las descripciones desempeñen, junto con la relación causal, un rol constitutivo del significado, esto es, en otros términos, que sean realmente introducidas en el mecanismo explicativo del significado. Si bien considero que esta objeción es fundamental, la dejaré provisionalmente de lado para centrarme en otros aspectos de la teoría que también encuentro discutibles, esto es, aspectos que podrían ser cuestionados aun bajo el presupuesto de que las descripciones asociadas fueran, en contra de la opinión de los autores, constitutivas del significado.

Devitt y Sterelny presentan dos explicaciones distintas, para distintos tipos de términos, del mecanismo de fijación de la referencia, afectado, como vimos, por el problema de la ambigüedad. Por un lado, sostienen que la referencia de los términos no básicos del lenguaje es fijada mediante la acción conjunta de la relación causal y el

uso de cierto tipo de descripciones, generales o categoriales; se trata, por consiguiente, de un mecanismo descriptivo-causal de fijación de la referencia. Por otro, la referencia de los términos básicos, es decir, las mencionadas descripciones categoriales así como los demostrativos simples, es fijada exclusivamente en virtud de la existencia de una relación causal, por lo que el mecanismo referencial en cuestión es puramente causal.[108] De esta manera, pretenden asimismo solucionar el antes mencionado problema de incompletitud que caracteriza a las teorías puramente descripcionales, puesto que hay algún punto -el nivel de las representaciones básicas- en el cual el lenguaje se conecta directamente con el mundo.[109] Ahora bien, en mi opinión, las dos tesis acerca de la fijación de la referencia aquí involucradas presentan algunos problemas.

Por un lado, la afirmación de que el único tipo de descripciones que juega un rol semánticamente pertinente en el bautismo inicial de individuos está constituido por descripciones en términos generales o categoriales no parece tener fundamento empírico alguno. En particular, tampoco parece estar fundada en evidencia psicológica ni psicolingüística: si se toma en cuenta la evidencia en cuestión, es plausible afirmar que tanto las descripciones indefinidas (como, por ejemplo, "una estructura") como las descripciones definidas (tales como "mi gato") pueden jugar un rol determinante en el bautismo inicial de individuos; más aun, hay ocasiones en que es sólo este tipo de descripciones y no las descripciones en términos categoriales las que juegan el rol en cuestión. A modo de ejemplo, cuando se bautiza a un recién nacido, se tienen sin duda en mente descripciones indefinidas tales como "un bebé precioso" y descripciones definidas tales como "mi hijo" y no veo razón para no creer que sean estos tipos de descripciones y no las categoriales mencionadas por Devitt y Sterelny, tales como "un individuo" o "la causa de mi percepción",

108. Cabe aclarar que hay un conjunto amplio de representaciones, los términos de clases artificiales, para los que Devitt y Sterelny proponen, a diferencia de Putnam, una explicación *puramente descripcional*. Véase Devitt & Sterelny 1987: pp. 75-79.
109. Véase Devitt & Sterelny 1987: p. 75.

las que contribuyen a la determinación del referente. Es más, considero que no hay evidencia clara de que éste último tipo de descripciones esté presente en casos como el del ejemplo.

Considero que es sólo para evitar ciertos problemas y esbozar la solución de otros por lo que Devitt y Sterelny se concentran exclusivamente en las descripciones categoriales. En otras palabras, la asignación de un rol clave a las descripciones en cuestión, lejos de basarse en evidencia apropiada, parece depender de decisiones *ad hoc*, relacionadas en parte con el "diseño arquitectónico" de la teoría. Mis razones en apoyo de esta tesis son las siguientes. A primera vista, parecería que la introducción exclusiva de descripciones de tipo categorial permite reducir a su mínima expresión el problema de la ignorancia y el error: lo que requieren del hablante no es conocimiento de propiedades específicas sino tan sólo conocimiento de tipo general -que es incluso admitido por el mismo Kripke. Sin embargo, considero erróneo pensar que esto es así: requerir el conocimiento de propiedades específicas no tiene porqué representar un requisito epistémico más fuerte que requerir el conocimiento de propiedades categoriales, puesto que se trata en ambos casos de conocimiento. Lo importante, y lo que distingue a este tipo de teorías de las teorías puramente descripcionales, es que la identificación del referente no depende exclusivamente del conocimiento en cuestión sino conjuntamente de éste y de la existencia de una relación causal.

Además, parecería que Devitt y Sterelny tienen en mente el desarrollo completo de un programa semántico que consiste en la propuesta de explicaciones descripcionales y mixtas para la mayor parte de los términos del lenguaje y explicaciones puramente causales para ciertos términos básicos; de este modo, la tarea semántica debe ajustarse a los dos requisitos siguientes: por un lado, es preciso ofrecer una explicación puramente causal de los términos básicos; por otro, se debe ofrecer también una explicación de los términos no básicos en términos de los términos básicos y de relaciones causales. Y aquí es donde entran en juego las descripciones categoriales: éstas parecen ser un candidato más plausible que las descripciones comunes para cumplir (junto con los demostrativos puros) el rol de los términos básicos. Si esto es así, dada la evidencia consignada más arriba,

la asignación de un rol clave a las descripciones categoriales parece basarse exclusivamente en razones de buen diseño arquitectónico de la teoría. Pero, en la medida en que el buen diseño arquitectónico de una teoría no aumenta necesariamente su plausibilidad, el rol asignado a las descripciones categoriales por Devitt y Sterelny es decididamente cuestionable.

En lo que respecta al bautismo inicial de las clases naturales, los autores no son claros acerca de qué tipo de descripciones deben ser introducidas. Según sus propias palabras, citadas anteriormente, el fundador de un término de clase natural debe asociar con el término en cuestión, en primer lugar, "alguna descripción que en efecto clasifique a ese término como un término de clase natural" y, en segundo lugar, "algunas descripciones que determinen qué naturaleza de la muestra es pertinente para la referencia del término". Respecto del primer punto, parecería tratarse de una descripción de tipo categorial ("muestra o ejemplar de una clase natural" en oposición a "individuo"); pero, nuevamente, ¿está fundada esta consideración en algún tipo de evidencia psicológica o psicolingüística? En otras palabras, ¿fue acaso el uso de ese tipo de descripciones lo que permitió, por ejemplo, llamar "agua" al agua y "fuego" al fuego? Dudo que este punto pueda ser establecido con certeza. En lo que concierne al segundo punto, no queda en absoluto claro de qué tipo de descripciones se trata. Los mismos Devitt y Sterelny son conscientes de este problema:

> *Debe haber algo acerca del estado mental del fundador que determine qué naturaleza putativa de la muestra es pertinente para el bautismo, de modo tal que si la muestra en cuestión no tiene tal naturaleza la fundación fracasa. Es muy difícil decir exactamente qué determina la naturaleza pertinente.*[110]

En principio, no veo obstáculo alguno para considerar que se trata de descripciones indefinidas y definidas, de distinto grado de especificidad: es más, ésta parece ser la única manera de poder distinguir

110. Devitt & Sterelny 1987: pp. 73-4.

una clase natural de otra. A modo de ejemplo, es plausible pensar que quienes bautizan una especie tal como la especie Hombre u Homo Sapiens han de tener en mente descripciones específicas tales como "el tipo de ser cuyos restos fueron hallados en Africa Central", "el sucesor del Homo Erectus", "el tipo de ser cuyo cráneo mide X centímetros", etc: descripciones generales tales como "animal", "vertebrado", "mamífero" sin duda no bastan para distinguir la especie Homo Sapiens de otras especies semejantes.

Por otro lado, como se mencionó anteriormente, Devitt y Sterelny consideran que es posible delimitar un subconjunto de representaciones básicas, a las cuales se aplique una explicación semántica puramente causal. De este modo, se presupone que la semántica primaria, esto es, la semántica de las representaciones básicas, puede ser desarrollada en términos (puramente) histórico-causales. Ahora bien, en mi opinión, no hay ningún subconjunto de representaciones básicas cuyas relaciones histórico-causales con los objetos del mundo sean unívocas; en otras palabras, dado el problema del respecto, no creo que sea posible especificar, como sostienen Devitt y Sterelny, la referencia de ciertas representaciones básicas en términos puramente histórico-causales. Ante todo, cabe destacar que los autores no son completamente claros y precisos respecto a las representaciones que formarían parte del subconjunto en cuestión: sólo sugieren que tal subconjunto podría estar integrado por demostrativos simples o puros y por términos y descripciones categoriales.

En lo que respecta a los demostrativos simples o puros, es decir, aquellas expresiones demostrativas que carecen de todo componente descriptivo, tales como "esto", "eso" y "aquello", cabe señalar que no voy a profundizar en este tema por considerarlo especialmente complejo. Baste señalar, sin embargo, que dudo que pueda darse una explicación puramente causal de su significado: sus conexiones causales directas con el mundo son indudablemente múltiples.[111]

111. Nótese que la manera en que la teoría desarrollada por Kaplan logra evitar este problema es el reconocimiento, junto con el referente, de otro factor que, según él, no es parte del significado sino precondición o requisito del mismo, denominado "carácter". Como puede apreciarse, esto constituye otro ejemplo de la antes mencionada estrategia de exportación.

En otras palabras, la teoría histórico-causal no parece contar con recurso alguno que permita responder a la crítica de Wittgenstein a las definiciones ostensivas, es decir, aquéllas que hacen un uso esencial de los demostrativos puros:

> *(...)la definición ostensiva sola no basta, porque cabe siempre la posibilidad de interpretarla de varias maneras. Por ejemplo, supongamos que yo explico la palabra 'tove' señalando un lápiz y diciendo 'Esto se llama tove'. La explicación sería completamente inadecuada, porque se puede pensar que significa 'Esto es un lápiz', o 'Esto es cilíndrico', o 'Esto es madera', o 'Esto es uno', o 'Esto es duro', etc.*[112]

Centrémonos entonces en el otro candidato sugerido, los términos y descripciones categoriales. Los ejemplos paradigmáticos están constituidos por las descripciones "individuo completo" y "miembro de una clase natural": ahora bien, no veo de qué manera tales descripciones pueden referir, sobre la sola base de una relación causal, a un objeto unívocamente en tanto individuo completo o miembro de una clase natural. Lo mismo puede decirse de otro de los ejemplos de categorías propuesto, "causa de O", donde "O" se considera que refiere al conjunto de las características observadas que son esenciales al objeto: la sola relación causal no parece ser suficiente para hacer que "causa de O" quede fundado en un objeto, de manera unívoca, en tanto causa de O. En un artículo inédito, Devitt sostiene que la relación causal involucrada debe ser entendida como una correlación confiable "en el momento de la interacción", la cual puede no ser mantenida en el futuro; sin embargo, más allá de que esto parece introducir un elemento, la confiabilidad, que es ajeno al modelo histórico-causal, no veo cuál puede ser el fundamento de una correlación causal de ese tipo.[113]

112. La cita está extraída de Kenny 1973: p. 141, pero constituye a su vez una cita de la *Gramática filosófica* de Wittgenstein.
113. Véase Devitt (inédito).

El mismo Devitt parece considerar que este punto constituye uno de los puntos más débiles de la teoría histórico-causal, como ponen de manifiesto las siguientes palabras:

> *No pretendo que la teoría histórico-causal de la referencia ejemplificada por IT [illustrative theory] sea completamente adecuada ni pretendo demostrar que yo sé cómo transformarla en una teoría tal. Pienso que tiene un problema profundo, 'el problema del respecto' (...), el cual dudo que pueda ser resuelto.*[114]

Sobre la base de lo anterior, considero que la teoría descriptivo-histórico-causal de la referencia, si bien parece ofrecer una explicación plausible del significado de gran parte de las representaciones de nuestro lenguaje y/o pensamiento, no provee los recursos suficientes para explicar el significado de ciertas representaciones básicas -cuyo dominio parece requerirse para la comprensión del resto; en otros términos, la semántica primaria queda claramente fuera de su alcance. Esto no implica que sea preciso abandonar la explicación propuesta para el significado de las representaciones no básicas: en mi opinión, la teoría descriptivo-histórico-causal puede ser preservada bajo la condición de que sea complementada mediante otro tipo, no histórico, de teoría causal. Como se vio en este capítulo, la ansiada explicación puramente causal de las representaciones básicas no puede darse en términos de relaciones histórico-causales pero tal vez pueda darse en términos de relaciones causales no históricas. En el capítulo siguiente, me ocupo entonces del examen de ciertas teorías recientes que, sin abandonar el modelo causal, proponen una explicación distinta, en un sentido ahistórica, de la aptitud referencial de nuestras representaciones. El propósito final de este examen es determinar cuál de ellas debe ser adoptada de modo tal de complementar la explicación descriptivo-causal propuesta en el presente capítulo.

114. Devitt 1996: p. 169. El destacado es mío.

Capítulo 3
Las teorías causales no históricas

En este capítulo, me ocupo de las teorías causales no históricas, es decir, de aquellas teorías puramente causales que no pertenecen a la tradición inaugurada por Kripke. Estas teorías se dividen en dos grupos principales. Por un lado, están las teorías indicativo-causales o informacionales, que definen la referencia y el significado en términos de información, la cual es a su vez definida en términos de correlaciones confiables entre las representaciones y los objetos del mundo. Por otro lado, están las teorías teleológico-causales o biosemánticas, que lo hacen en términos de la noción de función propia biológica o *télos* de las representaciones. Para realizar el análisis y evaluación de estas teorías, he decidido seleccionar a una teoría representativa de cada uno de los grupos anteriores. Más específicamente, en primer lugar, me ocupo de la teoría covariacional propuesta por Fodor, que constituye un ejemplo paradigmático del enfoque indicativo-causal, para luego centrarme en la teoría de Millikan, esto es, el ejemplo más representativo de la concepción biosemántica. El objetivo central de este examen ha sido, como señalé al final del capítulo anterior, establecer cuál de ellas resulta más adecuada para desempeñar el rol de semántica primaria, es decir, para ofrecer una explicación del significado de las representaciones básicas y constituir

de este modo el complemento requerido de la teoría descriptivo-histórico-causal antes propuesta. Es oportuno adelantar que la conclusión obtenida a partir del análisis y, por tanto, la tesis defendida en este capítulo es que el complemento en cuestión debe darse en términos de una teoría teleológica, en la línea general propuesta por Millikan. Más específicamente, considero que las explicaciones indicativo-causales, del tipo de la ofrecida por Fodor, no son adecuadas, por varias razones que se detallan a continuación, todas ellas relacionadas en última instancia con el excesivo sesgo verificacionista de este tipo de enfoque. Por otro lado, considero que la idea central del enfoque teleológico, la de función propia biológica, es lo suficientemente rica como para dar lugar al desarrollo de una visión adecuada de cómo nuestras representaciones básicas adquieren el significado que tienen. En consecuencia, sugiero que es plausible defender una combinación de la teoría descriptivo-causal presentada en el capítulo anterior para gran parte de las representaciones no básicas y una teoría teleológica tal como la propuesta por Millikan para el subconjunto de representaciones básicas; la combinación sugerida presupone una concepción de la relación entre el pensamiento y el lenguaje, que es explicitada en la conclusión del libro.[115]

Para concluir con esta breve introducción, quisiera hacer dos aclaraciones, una referida al estilo expositivo y la otra de carácter conceptual. En primer lugar, la tesis positiva defendida ha de ser reconstruida a partir de las críticas ofrecidas a otras teorías: nuevamente, es con ocasión de la crítica de lo que considero inadecuado que intento presentar mi propia visión de la cuestión estudiada. En segundo lugar, la tesis positiva defendida es, como señalé anteriormente, una tesis programática, esto es, defiendo un programa de trabajo semántico cuyos múltiples detalles requieren un grado de elaboración que excede los límites de este libro.

115. Cabe destacar que una combinación de este tipo, pero basada en distintos argumentos, ha sido propuesta, como mencioné en la introducción, por Sterelny en Sterelny 1990 (véase especialmente el capítulo 6).

1. La teoría covariacional de Fodor

1.1. ¿Es fatal la ambigüedad?

Como señalé anteriormente, la noción de causalidad parece ofrecer a la semántica una vía de escape del círculo intencional: si la referencia puede ser explicada en términos causales, el ideal fisicalista de reducir todos los conceptos, incluidos los conceptos semánticos, a conceptos físicos deja de ser una meta inalcanzable. Sin embargo, como muestran los dos capítulos anteriores, el proyecto está ensombrecido: la causalidad no resulta ser un concepto lo suficientemente poderoso como para desempeñar el rol deseado. De este modo, las teorías histórico-causales de la referencia deben incorporar factores descriptivos y transformarse en teorías mixtas. Aun así, es preciso disponer de una explicación puramente causal para el subconjunto de representaciones básicas o, en otros términos, de una semántica primaria puramente causal, a riesgo de dejar sin explicar la conexión última entre el sistema representacional en su conjunto y el mundo, es decir, de proporcionar una explicación circular y por ello fundamentalmente incompleta del significado. Ahora bien, como vimos al final del capítulo anterior, el problema de ambigüedad denominado "problema del respecto" ("*qua problem*") se resiste firmemente a obtener una solución definitiva en el marco de la teoría mixta más promisoria, la teoría descriptivo(-histórico-)causal. Es esto justamente lo que conduce al planteo de la pregunta inicial: ¿es, acaso, fatal la ambigüedad? Dicho de otra manera, ¿involucra acaso el engorroso círculo antes mencionado una objeción para toda explicación dada en términos causales?

En este punto, es oportuno recordar que, además de explicaciones histórico-causales, se han propuesto también otro tipo de explicaciones causales, tales como las indicativo-causales o informacionales. Consideremos entonces la propuesta de las teorías indicativo-causales, cuyo origen puede situarse en la obra de Dretske.[116] El dominio de

116. Véase Dretske 1981, especialmente la parte III.

este segundo tipo de teorías suele ser el pensamiento o lenguaje mental. En general, todas ellas sostienen que el significado puede ser explicado en términos de información: afirmar que los conceptos tienen significado equivale a afirmar que éstos nos informan acerca del mundo, del mismo modo en que los anillos de un árbol nos informan acerca de su edad o que el humo nos informa acerca de la presencia de fuego; en otras palabras, los conceptos son índices o indicadores naturales. En términos de Dretske:

> *Un concepto primitivo requiere que el sistema que lo posee tenga la capacidad para recibir, y de hecho haber recibido, información correspondiente al significado de ese concepto (su contenido semántico).*[117]

Ahora bien, las relaciones de información involucradas deben entenderse en términos de relaciones causales confiables. En general, un concepto refiere a un cierto objeto si y sólo si contiene información acerca de él, y esto último ocurre si y sólo si el concepto está confiablemente correlacionado con el objeto en cuestión, esto es, si covaría con aquél.

Una relación causal es considerada confiable o de covariancia cuando es legal o nómica; lo que fundamenta la existencia de relaciones legales o nómicas entre conceptos y objetos son nuestras capacidades discriminatorias, esto es, nuestras capacidades para distinguir unos objetos de otros -lo que hace que cada objeto esté confiablemente correlacionado o covaríe con la instanciación de un determinado concepto y no de cualquier otro. Como es sabido, el carácter legal de una relación causal es garantizado por la verdad de una oración contrafáctica -de ahí que se suela caracterizar a tales relaciones como "relaciones que soportan contrafácticos" ("*counterfactual supporting*"). A modo de ejemplo, dado que es verdadero, en virtud de ciertas propiedades físicas del mundo, que en condiciones similares si la temperatura no descendiera a 0° el agua pura no se congelaría, es posible

117. Dretske 1981: p. 223.

afirmar que el descenso de la temperatura a 0° causa confiablemente el congelamiento del agua pura; del mismo modo, desde el punto de vista indicativo-causal, dado que es verdadero, en virtud de la naturaleza de nuestras capacidades discriminatorias, que en condiciones similares si no existiera la rojez (o la propiedad de ser rojo o la clase de las cosas rojas) no existiría el concepto ROJO, es posible afirmar que la rojez causa confiablemente casos de ROJO -y, por tanto, que ROJO contiene información acerca de la rojez, o, en otras palabras, que ROJO significa rojez.[118] De acuerdo con este tipo de teorías, el significado de un concepto está entonces constituido no meramente por aquello que de hecho lo causa sino fundamentalmente por lo que lo causaría, en condiciones similares, en cualquier situación posible: éste es exactamente el concepto de causa confiable. De este modo, mientras que las teorías histórico-causales apelan sólo a las relaciones causales que de hecho ocurren entre las representaciones y los objetos, las informacionales o indicativo-causales consideran no sólo las relaciones efectivas sino también aquéllas que podrían haberse dado si el mundo hubiera sido levemente distinto; en síntesis, no sólo apelan a relaciones causales efectivas sino también a aquéllas meramente posibles o contrafácticas.

¿Cómo resuelve una teoría de este tipo el problema del respecto? En términos del ejemplo anterior, ¿sobre qué base nos permite afirmar que ROJO refiere a la rojez y no a una cierta intensidad de rojo? Como vimos, para determinar la referencia de ROJO, es preciso tener en cuenta no solamente las relaciones causales efectivas, que involucran tanto a la rojez como a una cierta intensidad de rojo, entre otras cosas, sino también ciertas posibilidades contrafácticas. En condiciones similares, si se nos hubiera presentado algún otro tono de rojo en lugar del que de hecho se nos presenta, habríamos formado igualmente el concepto ROJO. Por consiguiente, no es la relación causal entre una cierta intensidad de rojo y un caso de ROJO

118. Vale la pena destacar que el requisito de que se preserven las condiciones iniciales es un factor muy importante en la teoría -establecido por medio de las denominadas "cláusulas *ceteris paribus*".

lo que está subsumido por una ley. En otras palabras, es nuestra capacidad para distinguir la rojez de otras propiedades, y no nuestra capacidad para distinguir un tono de rojo de otro, lo que fundamenta la covariancia nómica entre ROJO y una cierta propiedad del mundo. Dada nuestra peculiar capacidad discriminatoria, podemos estar seguros de que el significado o contenido de ROJO no es una cierta intensidad de rojo sino la rojez en tanto *-qua-* rojez.[119]

Sin embargo, como podrá apreciarse, esta propuesta no logra impedir que reaparezca nuestro viejo problema con un nuevo rostro. Nuestras capacidades discriminatorias no son en absoluto estables: permiten, para poner un ejemplo de Fodor, que los caballos en la oscuridad causen casos de VACA; en términos de ese autor, permiten instancias salvajes de VACA.[120] ¿Sobre qué base es entonces posible seleccionar la relación causal vaca-VACA como aquélla que determina el significado de VACA o, desde la perspectiva conversa, sobre qué base es posible descartar a la relación causal caballo en la oscuridad-VACA como errónea? En otras palabras, dado que tanto las vacas como los caballos en la oscuridad y muchas otras cosas más que, debido a la falibilidad de nuestras capacidades discriminatorias, pueden ser confundidas con vacas causan confiablemente casos de VACA, ¿cómo es posible negar que el significado de "vaca" sea el significado disyuntivo <vaca o caballo en la oscuridad o...>?[121] Además, para poner otro ejemplo, muchas instancias de AGUA parecen ser confiablemente causadas por comida salada; ¿cuál es el fundamento para eliminar la posibilidad de que AGUA signifique <comida salada> o algo semejante? En términos generales, dado que el significado ha sido definido en términos de la noción de causa confiable,

119. Cabe aclarar que, a lo largo de este capítulo, dado que las teorías analizadas se refieren originariamente (y, en el caso, de la teoría covariacional, de manera exclusiva) al lenguaje del pensamiento, la expresión "contenido" es utilizada con mayor frecuencia que la expresión "significado". Recuérdese que, si bien son intercambiables en casi todos los contextos, en la discusión en torno a las teorías psicosemánticas, es más frecuente el uso de "contenido".

120. Véase Fodor 1987: cap. 4.

121. Cabe señalar que los corchetes angulares son la convención elegida para designar significados, como quiera que se los conciba.

se requiere una explicación de cómo es posible que el significado de un signo sea insensible a las diversas causas confiables de sus distintos casos o instancias, a los diversos objetos del mundo acerca de los cuales acarrea información: sin una adecuada explicación de la llamada "robustez" del significado, las teorías indicativo-causales no pueden sostenerse. En términos de Fodor:

> *La información está unida a la etiología de una manera en la cual el significado no lo está. Si los casos de un símbolo tienen dos tipos de etiologías, se sigue que hay dos tipos de información acarreada por los casos de ese símbolo. (Si algunos casos de 'vaca' son causados por vacas y algunos otros casos no lo son, se sigue entonces que algunos casos de 'vaca' acarrean información acerca de las vacas y algunos otros no lo hacen.) En contraste, el significado de un símbolo es una de las cosas que todos sus casos tienen en común, comoquiera que hayan sido causados. Todos los casos de 'vaca' significan vaca; si así no fuera, no serían casos de 'vaca'.*[122]

Como puede apreciarse, se trata, al igual que en el caso del problema del respecto, de un problema de ambigüedad, conocido como "problema del error o la disyunción". La diferencia entre uno y otro es la siguiente: en el caso del problema del respecto, la ambigüedad está determinada por la imposibilidad de seleccionar uno entre los múltiples aspectos de un objeto históricamente relacionado con una cierta representación; lo que da lugar a la ambigüedad en el caso del problema del error o la disyunción es, en cambio, la imposibilidad de seleccionar uno entre los múltiples objetos confiablemente correlacionados con una representación dada. Sterelny caracteriza al primero como un problema de profundidad (*depth problem*), concerniente a la delimitación de qué es lo que cuenta como objeto referido, mientras que el segundo es un problema de ancho o extensión (*breadth problem*), concerniente a la delimitación de cómo ese objeto puede ser extendido.[123]

122. Fodor 1990: p. 90.
123. Sterelny 1990: p. 113.

En síntesis, las explicaciones causales son fuertemente atractivas: descansan sobre una noción, la de causalidad, que promete ofrecer una explicación de la intencionalidad acorde con el ideal fisicalista. Sin embargo, sus distintas versiones presentan problemas similares. Por un lado, las teorías histórico-causales no pueden solucionar el problema del respecto sin caer en un círculo que amenaza con despojarlas de todo poder explicativo. Por otro, las teorías indicativo-causales se enfrentan con el mencionado problema del error o la disyunción. Ahora bien, es el marco de éstas últimas el que ha elegido Fodor para proponer su teoría semántica, a saber, la teoría covariacional. De este modo, la teoría de Fodor representa fundamentalmente un intento por solucionar el problema de la disyunción. En lo que sigue, me centraré en esta propuesta.

1.2. La respuesta de Fodor

De acuerdo con lo anterior, la pregunta fundamental que Fodor se propone responder es la siguiente: ¿cómo puede el significado, definido en términos de información, ser indiferente a ésta última? Para retomar el ejemplo anterior, dado que tanto las vacas como los caballos en la oscuridad causan confiablemente nuestras representaciones de vaca, y por tanto debemos considerar que éstas contienen información tanto acerca de las vacas como de los caballos en la oscuridad, ¿cuál es el fundamento para afirmar que VACA significa <vaca> y no <caballo en la oscuridad> ni <vaca o caballo en la oscuridad>? En otros términos, ¿cuál es el fundamento para afirmar que la relación causal caballo en la oscuridad-VACA es errónea o que VACA no tiene un significado disyuntivo?

Según Fodor, y hasta aquí coincide con cualquier otra teoría informacional, a fin de contestar esta pregunta, hay que tener en cuenta no sólo lo que de hecho ocurre en el mundo tal como es sino también lo que habría ocurrido si el mundo hubiera sido levemente distinto; en otras palabras, debemos aplicar el llamado "método de las diferencias a través de los mundos posibles". El aporte específico de Fodor aparece en el paso siguiente: en términos del

ejemplo anterior, es necesario percatarse de que, en condiciones similares, si las vacas no hubieran causado casos de VACA, los caballos en la oscuridad tampoco los habrían causado. La relación causal entre los caballos en la oscuridad y los casos de VACA parece depender de la relación causal entre las vacas y los casos de VACA. Por otro lado, es preciso tener en cuenta que, en condiciones similares, si los caballos en la oscuridad no hubieran causado casos de VACA, las vacas los habrían causado de todas maneras. La relación causal entre las vacas y los casos de VACA no parece depender, en cambio, de la relación causal entre los caballos en la oscuridad y los casos de VACA. De ahí que sea posible concluir que la relación causal entre los caballos en la oscuridad y los casos de VACA depende asimétricamente de la relación causal entre las vacas y los casos de VACA. Según Fodor, la existencia de esta dependencia asimétrica muestra que la relación causal vaca-VACA es básica y determinante del significado. De este modo, se tiene un fundamento para afirmar que el significado de VACA no es un significado disyuntivo o ambiguo tal como <caballo en la oscuridad o vaca> sino que está unívocamente determinado. Lo que aquí subyace es una vieja intuición: los usos erróneos de un concepto dependen ontológicamente de sus usos verídicos, y no a la inversa. De esta manera, Fodor logra romper la simetría entre unos y otros.

A primera vista, ésta parece ser una ingeniosa solución a los problemas planteados por el fenómeno del significado y la intencionalidad. Desde esta perspectiva, los significados son explicados en términos de relaciones referenciales, y éstas son a su vez explicadas en términos de relaciones causales, cuya unívoca determinación puede llevarse a cabo mediante la aplicación del método de las diferencias a través de los mundos posibles. De este modo, que implica el abandono temporario del mundo actual y el consiguiente vuelo sobre el espectro de los mundos posibles, se hace posible distinguir los correlatos ontológicos de la relación causal que cumplen un rol semántico de aquéllos que no lo hacen. La determinación del significado de un signo no requiere entonces la consideración diacrónica de sus usos efectivos sino el examen sincrónico de sus posibles usos. La mirada histórica conduce, en última instancia, al

momento del bautismo inicial, signado por el problema del respecto. En contraste, la consideración de las posibilidades contrafácticas parece permitir la ansiada ruptura del engorroso círculo intencional. Creo, sin embargo, que la propuesta de Fodor presenta algunos problemas, que me propongo explicitar en la sección siguiente.

1.3. La crítica de la teoría covariacional

1.3.1. La cláusula ceteris paribus

Es importante destacar que con el término "capacidades discriminatorias" Fodor no pretende referirse a las capacidades discriminatorias individuales, ya sea como parte de una psicología *standard* o como expresión de peculiaridades psicológicas. Las capacidades discriminatorias individuales no tienen por qué fundamentar relaciones causales legales semánticamente pertinentes; por lo tanto, sería inadecuado pretender identificar a éstas últimas sobre la base del ejercicio de tales capacidades.[124] De acuerdo con esto, la teoría de Fodor apela a un procedimiento de idealización: las capacidades discriminatorias involucradas son las capacidades de discriminación de un observador ideal. Es entonces desde la perspectiva de quien no es ignorante ni comete errores ni realiza asociaciones arbitrarias que deben considerarse los mundos posibles. Esto involucra,

124. A modo de ejemplo, si bien yo no soy capaz de distinguir a los olmos del resto de los árboles, todas mis instancias de OLMO significan <olmo>; en otras palabras, las relaciones causales semánticamente pertinentes no parecen fundarse en mis capacidades discriminatorias personales: los olmos, para mí, indistinguibles de los demás árboles, nunca causan en mí casos de OLMO, y sin embargo, todas mis usos de OLMO significan <olmo>. ¿Por qué entonces la determinación del significado de OLMO habría de depender de mi capacidad para discriminar olmos en una situación contrafáctica? Dada la inutilidad semántica de mis capacidades discriminatorias en el curso real de mi vida, ¿por qué pensar que podrían resultar útiles en situaciones contrafácticas? No hay ninguna razón para creer que los significados dependan de tales rasgos subjetivos.

al mismo tiempo, la imposición de un límite a los mundos en cuestión: debe considerarse sólo a aquéllos que se parezcan lo más posible al mundo actual -más específicamente, que sólo difieran del mundo actual en la interpretación del concepto en estudio. En el ejemplo utilizado anteriormente, si se trata de determinar el significado de VACA, adoptar el punto de vista del observador ideal implica considerar sólo a aquellos mundos posibles que sean exactamente como el nuestro excepto o bien por el hecho de que en ellos sólo las vacas causan casos de VACA o bien por el hecho de que en ellos nada causa casos de VACA. De ahí que la idealización constituya el fundamento de la inclusión de la denominada "cláusula *ceteris paribus*" (expresión que es comúnmente traducida por el giro "en condiciones similares") antes mencionada: la adopción del punto de vista del observador ideal hace posible, mientras se determina el significado de un concepto, mantener constantes a todos los otros factores. Ahora bien, considero que lo que se dice acerca de las condiciones (ideales) en las que se supone que el mecanismo causal determina la relación semántica de referencia presenta algunos problemas.

En primer lugar, no hay un criterio claro que permita establecer, desde el punto de vista del observador ideal, cuáles son los mundos más parecidos o más cercanos al nuestro. Esto puede verse claramente en la respuesta dada por Fodor a una objeción hecha por Block que resumiré a continuación. Según Block, dado que VACA no es sino un signo codificado en el cerebro, ¿por qué no habría de ser posible pensar que puede ser causado por algo que no sea una vaca, por ejemplo, un árbol? En opinión de Block, no hay fundamentos suficientes para negar que haya mundos posibles en donde todos los casos de VACA son causados por caballos en la oscuridad.[125] Según la respuesta de Fodor, conceder la posibilidad planteada por Block equivale a pasar por alto un aspecto muy importante de las leyes que determinan el significado, a saber, el que señalamos en el parágrafo anterior: sus cláusulas *ceteris paribus*. En su opinión, si

125. Para una presentación más detallada de la objeción, véase Fodor 1990: pp. 111-112.

todo lo demás es similar al mundo actual, no es posible pensar que hay casos de VACA causados por caballos en la oscuridad sin que los haya causados por vacas. ¿Por qué? La razón es la siguiente:

> *Es intuitivo presuponer que los mundos que son exactamente como el nuestro con la sola excepción de que en ellos las vacas no causan casos de VACA son ipso facto más cercanos a nosotros que los mundos exactamente como el nuestro excepto por el hecho de que en ellos las vacas no causan casos de VACA y los árboles lo hacen.*[126]

En mi opinión, esto es muy oscuro. En la medida en que no se dé un criterio claro de cercanía entre mundos posibles, no podrá decidirse qué mundos son los más cercanos al nuestro y, por tanto, qué mundos deben ser tomados en cuenta para hacer las correspondientes discriminaciones. Uno puede pensar, contrariamente a lo señalado por Fodor, que cuando la relación causal x-X (por ejemplo, vaca-VACA) no se da, es justamente porque se da alguna otra y-X (por ejemplo, árbol-VACA). Imaginémonos, a modo de ejemplo, que la Tierra Gemela, descripta en el experimento mental de Putnam, es un mundo posible:[127] ¿está acaso la Tierra Gemela más alejada del mundo actual que otro mundo posible en donde no hay concepto alguno de agua -en otras palabras, en donde la relación causal H_2O-AGUA no se da y ninguna otra sustancia causa casos de AGUA? No lo creo, o, por lo menos, considero que no hay un criterio claro para responder este tipo de preguntas. De este modo, no es en absoluto claro cuáles son los contrafácticos que han de considerarse pertinentes para la determinación de la existencia de correlaciones confiables.

En segundo lugar, tengo la impresión de que hay cierta circularidad en la propuesta de Fodor. El pide que se considere, por un lado, (i) a aquellos mundos posibles, cercanos al actual, en donde las vacas causan casos de VACA pero ninguna otra cosa lo hace, en particular, no

126. Fodor 1990: p. 112.
127. Esto implica apartarse del experimento en cuestión, en donde se trata, por hipótesis, de un mundo *real* o *actual* -y no meramente posible.

lo hacen los caballos en la oscuridad; por otro, (ii) a aquellos mundos posibles, cercanos al actual, en donde nada causa casos de VACA. El objetivo, como vimos, es determinar si la relación vaca-VACA es o no una relación básica, de la que dependen las demás, como, por ejemplo, la relación caballo en la oscuridad-VACA. Ahora bien, por un lado, sobre la base de lo sugerido en el parágrafo anterior, no es claro que la consideración de (ii) involucre la consideración de un mundo cercano al actual: un mundo posible en donde se carece del concepto de vaca puede ser muy distinto y, por tanto, estar muy alejado del mundo actual. Por otro lado, lo que es más grave, si el observador ideal es capaz de cumplir con (i), no parece tener ningún problema por resolver. En otras palabras, el planteo de ese requisito por parte de Fodor parece presuponer aquello que se quiere concluir, a saber, que la relación básica, determinante de la relación semántica, es la relación vaca-VACA. Pero esto es justamente lo que cuestiona el problema de la disyunción: la relación básica no es vaca-VACA ni caballo en la oscuridad-VACA ni ninguna clase natural particular-VACA sino vaca o caballo en la oscuridad-VACA. El problema de la disyunción establece precisamente que la relación básica es la relación causal entre una clase no natural que incluye a las vacas, los caballos en la oscuridad, las cebras sucias de barro y todas aquellas cosas que puedan ser confundidas con las vacas (en definitiva, la clase no natural de las cosas que parecen vacas) y VACA. Esta es la relación causal básica porque es la más confiable: todos los usos de VACA son instancias de covariancia entre VACA y esa clase no natural. Si esto es así, considerar un mundo en donde las vacas causan casos de VACA es siempre considerar un mundo en donde la mencionada clase no natural causa casos de VACA, lo cual no equivale a considerar un mundo en donde sólo las vacas causan casos de VACA -que es lo que exige (i). En otras palabras, concebir la posibilidad de que sólo las vacas causen casos de VACA (y atribuir a un observador ideal el correspondiente poder discriminatorio, que está a la base de esa relación causal) implica pasar por alto el problema de la disyunción. Pero pasar por alto un problema es justamente lo contrario de solucionarlo.

En síntesis, la aplicación estricta del método fodoriano para determinar el significado (el denominado "método de las diferencias a

través de los mundos posibles"), en la medida en que se basa en la presuposición de que hay mundos observables desde un punto de vista ideal en donde sólo las vacas causan casos de VACA, parece trivializar el problema del significado. Si tales mundos son fácilmente discernibles, entonces el problema tiene una respuesta clara. Pero, en mi opinión, la presuposición involucrada no es razonable puesto que ignora el problema de la disyunción, que es lo que hace básicamente problemático al significado, desde la perspectiva indicativo-causal.[128]

1.3.2. El problema de Quine

Cabe destacar que es posible discernir un sentido en el cual la ambigüedad no parece ser disuelta mediante el recurso propuesto, a saber, la consideración de los mundos posibles cercanos al nuestro. En los términos del ejemplo anterior, bajo el supuesto de que sea posible identificar mundos cercanos al nuestro en donde las vacas, pero no los caballos en la oscuridad ni ningún otro objeto de apariencia semejante, causan casos de VACA, los mundos en cuestión serían sin duda mundos en donde también los estadios temporales de las vacas y las partes no separadas de las mismas causarían casos de VACA. De acuerdo con esto, VACA seguiría siendo ambigua, pues, en términos de Fodor, su contenido sería disyuntivo: <vaca o estadio temporal de vaca o ...o parte no separada de vaca>. Este tipo de ambigüedad, señalada por Quine en el desarrollo del argumento de la indeterminación de la traducción, podría ser denominada "ambigüedad quineana".[129]

Ante todo, cabe consignar que la ambigüedad quineana constituye un caso no del problema de la disyunción sino del problema del respecto, esto es, ilustra la tesis de que cada término está correlacionado con múltiples aspectos de un objeto. En términos del ejemplo de Quine, el conejo entero, la conejidad, los estadios

128. Una crítica en parte similar a la presentada en este punto se encuentra en Godfrey-Smith 1989.
129. Véase Quine 1960: cap. 2.

temporales de un conejo, sus partes no separadas, no son distintos objetos que pueden ser confundidos entre sí sino distintos aspectos bajo los cuales es posible considerar a un mismo objeto. De este modo, es preciso reconocer que, en oposición a lo que suele considerarse, no todas las teorías indicativo-causales ofrecen una solución al problema del respecto -o, por lo menos, no todas ellas permiten solucionar todas las instancias del mismo.

Ahora bien, considero que el análisis del problema de Quine conduce directamente al cuestionamiento del concepto central de la teoría, a saber, el de capacidad discriminatoria, el cual, no es debidamente especificado por Fodor. Hay elementos que indican, sin embargo, que alude a capacidades discriminatorias de tipo perceptual: el hecho, por ejemplo, de que Fodor plantee el problema de la disyunción como aquél surgido de la incapacidad para distinguir un objeto de todo otro que tenga su misma apariencia; si esto es un problema, lo que se requiere entonces, y para lo cual se plantea la necesidad del procedimiento de idealización, es la capacidad de distinguir correctamente a los objetos en virtud de sus apariencias (en términos de nuestro ejemplo, se requiere una situación en la cual si algo se me aparece como una vaca o es percibido por mí como tal sea en efecto una vaca y no otra cosa).[130] Ahora bien, si esto es así, el problema que quisiera plantear es el siguiente.

A primera vista, parecería que existe una clara oposición entre los casos de ambigüedad supuestamente resueltos por la teoría y los casos de ambigüedad quineana. En términos de ejemplos ya utilizados, la diferencia entre las vacas y los caballos en la oscuridad, así como la diferencia entre la rojez en general y los distintos matices de rojo es una diferencia que puede ser percibida; por el contrario, la diferencia

130. Véanse las siguientes palabras de Sterelny, en coincidencia con esta evaluación de la teoría: "Es difícil extender a las teorías indicativas de las explicaciones de la representación perceptual a las de la representación mental en general, y de los conceptos de tipos definibles en forma ostensiva a los otros conceptos." Sterelny 1990: p. 115. Y más adelante: "La teoría indicativa es, como mucho, una teoría de la representación perceptual; necesita ser extendida para abarcar a la representación cognitiva en general." p. 119.

entre los conejos y los estadios temporales de conejos no puede serlo. De ahí que, por un lado, (i) en tanto es posible, en condiciones ideales, distinguir perceptualmente a una vaca de cualquier otra cosa parecida, como un caballo en la oscuridad, pueda en definitiva decirse que VACA significa <vaca> y no <vaca o caballo en la oscuridad>; del mismo modo, en tanto es posible, en condiciones ideales, distinguir perceptualmente a la rojez de cualquier matiz particular de rojo, puede asimismo decirse que ROJO significa <rojez> y no <rojez o matiz particular de rojo>; por otro lado, en cambio, (ii) en tanto no es posible distinguir perceptualmente a un conejo de, por ejemplo, uno de sus estadios temporales, en definitiva, lo único que la teoría permite afirmar acerca de CONEJO es que significa <conejo o estadio temporal de conejo>. De acuerdo con esto, la razón por la que la teoría no puede resolver la ambigüedad quineana residiría en que su dependencia crucial del concepto de discriminación perceptual la torna inepta para lidiar con aquellos casos, como los planteados por Quine, en donde la capacidad de discriminación requerida no es perceptual sino que involucra la introducción de categorías semánticas y ontológicas -tales como las de términos de clases naturales, objetos materiales, etc. En términos del ejemplo inicial, la razón por la que aquellos mundos cercanos al nuestro en donde las vacas causan casos de VACA no pueden ser distinguidos de aquellos otros en donde los estadios temporales de las vacas causan casos de VACA está dada por el hecho de que, en condiciones ideales, no es posible discriminar perceptualmente a las vacas de sus estadios temporales: la distinción en cuestión requiere la intervención de otras partes de nuestro sistema conceptual, esto es, aquéllas que fundamentan la capacidad para discriminar la categoría de objetos enteros de la de estadios temporales de objetos.

En mi opinión, sin embargo, la oposición entre los casos supuestamente resueltos por la teoría y los ejemplos quineanos de ambigüedad no es tan clara. Como es sabido, la percepción está profundamente relacionada con otros capacidades cognitivas y no cognitivas. Si esto es así, no es claro que la discriminación entre vacas y caballos en la oscuridad pueda hacerse exclusivamente en términos de una capacidad perceptual: lejos de ello, parece requerir la intervención de cierta capacidad clasificatoria, ejercida en función de una cierta concepción del mundo -según la

cual éste está constituido no sólo por objetos enteros y no por estadios temporales de objetos sino también por vacas y caballos y no, por ejemplo, por animales grandes de cuatro patas que agradan a los dioses.

Pero, la apelación a capacidades discriminatorias no perceptuales parece encerrar a la teoría en un engorroso círculo: si se quiere explicar en términos compatibles con el fisicalismo de qué manera adquieren significado nuestros conceptos, no parece admisible utilizar en la explicación a esos mismos conceptos, con sus significados habituales. Dicho de otro modo, en términos del ejemplo anterior, si la explicación de porqué VACA significa <vaca> y no <estadio temporal de vaca> ni <caballo en la oscuridad> involucra la apelación al hecho de que el concepto VACA (con su significado) es parte de nuestro sistema conceptual, no parece ser realmente explicativa -esto es, aclaratoria o iluminadora del fenómeno involucrado. Desde cierta perspectiva, podría considerarse que se trata de un pedido de explicación superfluo: ¿qué sentido tiene preguntar algo cuya respuesta ya se conoce?

En definitiva, considero que el análisis del denominado "problema de Quine" conduce a cuestionar un elemento clave de la teoría, el concepto de capacidad discriminatoria. La principal objeción es que se trata de un concepto que no está lo suficientemente especificado. Como consecuencia de ello, se presentan dos opciones: o bien se presupone que alude a capacidades discriminatorias perceptuales, en cuyo caso el problema de Quine queda sin respuesta; o bien se interpreta que las capacidades discriminatorias cuya ejercicio se requiere para la determinación del significado son capacidades discriminatorias generales: en este segundo caso, es posible considerar que la teoría de Fodor soluciona la ambigüedad quineana pero a riesgo de involucrar un círculo conceptual que es, por lo menos, delicado.

1.3.3. Discriminación, idealización y verificacionismo

Si bien Fodor considera que la referencia está constituida por una relación causal directa entre conceptos y objetos, su explicación de cómo se determina esa relación involucra, como se ha visto, la apelación al ejercicio de capacidades discriminatorias en condiciones

ideales. Como se mencionó anteriormente, la puesta en marcha del mecanismo referencial depende del ejercicio idealizado de nuestras capacidades discriminatorias, puesto que sólo en ese contexto las correlaciones confiables adquieren pertinencia semántica. De este modo, se introduce claramente un factor epistémico en la teoría causal; en otras palabras, hay, en la teoría, un factor claramente epistémico que condiciona la operación del mecanismo causal. Considero que no puede decirse entonces, como pretende Fodor, que se trata de una teoría *puramente* causal de la referencia. Según lo que hemos visto hasta ahora, las teorías acerca de la referencia podían ser clasificadas en tres grupos principales: por un lado, aquéllas según las cuales una palabra o un concepto refiere a un objeto en virtud de una descripción del objeto referido asociada con la representación en cuestión; por otro, aquéllas que sostienen que una palabra o un concepto refiere a un objeto en virtud de una relación causal directa con el mismo, esto es, sin que el hablante se halle en posesión de ninguna descripción del objeto referido; finalmente, aquéllas que intentan combinar de alguna manera ambos mecanismos referenciales. La complejidad de la teoría de Fodor reside, en mi opinión, no sólo en el hecho de que no puede ser claramente ubicada en ninguno de estos grupos sino también en el hecho de que la combinación propuesta es difícil de desentrañar. Propongo entonces detenerse en el análisis del concepto clave de capacidad discriminatoria idealizada involucrado por la teoría.

En primer lugar, y como se mencionó anteriormente, no queda claro a qué se refiere Fodor con la expresión "capacidad discriminatoria". Dada la falta de especificaciones al respecto, es plausible interpretar esa expresión como haciendo referencia a una de las formas que puede adoptar el conocimiento identificador -cabe destacar que, según el sentido común del término, discriminar algo presupone ser capaz de identificarlo o diferenciarlo de otras cosas. Más aun, si, como es posible interpretar, discriminar algo implica percibirlo, entonces la capacidad epistémica requerida al observador ideal por la teoría de Fodor involucra una exacerbación de la capacidad epistémica requerida a los hablantes comunes por la concepción descripcional -dado que la capacidad

de percepción representa un requisito epistémico más estricto que la capacidad de descripción.

En segundo lugar, cabría pensar que el peso de la objeción anterior es significativamente aligerado por la necesidad de tomar en cuenta que el conocimiento identificador no es requerido del hablante individual sino del hablante *ideal*. En otras palabras, parecería que, en virtud de la idealización, el peso del factor epistémico en la teoría de Fodor es menor que en las teorías descripcionales. Sin embargo, es lícito preguntarse cuál es, exactamente, la relación entre las capacidades discriminatorias de los hablantes individuales y las del hablante ideal. Pues bien, el poder discriminatorio del hablante ideal no es sino el que tendría el hablante común si no tuviera las limitaciones físicas y psicológicas que le son propias; en algún sentido, podría decirse, sin embargo, que se trata de las mismas capacidades discriminatorias: lo que varía, en cada caso, son las condiciones para su ejercicio. En términos de la distinción entre competencia y actuación, podría decirse que los hablantes individuales son tan competentes como el hablante ideal: lo que distingue a los unos del otro es la respectiva actuación.[131] Lo que pretendo destacar con este punto es que, en definitiva, la teoría apela de alguna manera al poder de discriminación del hablante individual, en tanto apela a un poder de discriminación del mismo tipo, que constituye su modelo -un ejemplo acabado de cómo debería ser ejercido. De este modo, los objetos de referencia resultan ser, en algún sentido, objetos discriminables, a los que el hablante común tiene la capacidad de identificar -aunque no la desarrolle plenamente debido a sus múlti-

131. La distinción entre los conceptos de competencia y actuación fue formulada por Chomsky en el marco de su teoría sintáctica de los lenguajes naturales. Véase, por ejemplo, el párrafo siguiente: "Hacemos, pues, una distinción fundamental entre COMPETENCIA (el conocimiento que el hablante-oyente tiene de su lengua) y ACTUACIÓN (el uso real de la lengua en situaciones concretas). Sólo en la idealización establecida en el párrafo anterior es la actuación reflejo directo de la competencia. En la realidad de los hechos, es obvio que no puede reflejar directamente la competencia. Cualquier testimonio del habla natural mostrará numerosos arranques en falso, desviaciones de las reglas, cambios de plan a mitad de camino y demás". Chomsky 1970: p. 6.

ples limitaciones naturales.[132] Bajo esta interpretación, es indudable que el factor epistémico tiene un peso considerable en la teoría, mayor de lo que podría pensarse inicialmente.

En este punto, es oportuno recordar entonces las objeciones kripkeanas basadas en la ignorancia y el error mencionadas en el primer capítulo. Es pertinente aclarar que el argumento de Kripke puede servir para defender no sólo la postura un poco extrema y tal vez insostenible, según la cual la referencia es compatible con ignorancia total y error completo acerca de las propiedades del referente, sino también la postura, más moderada, según la cual la referencia es compatible con ausencia de conocimiento que permita la identificación unívoca del objeto referido, entre cuyas formas puede incluirse al conocimiento discriminatorio. De este modo, si es cierto que por lo menos cierto grado de ignorancia y error no son obstáculos para la referencia, ¿por qué postular una idealización que los evite? En la medida en que el mecanismo explicativo conlleve una idealización de ese tipo, no parece ser descriptivo de la práctica lingüística cotidiana -lo cual representa una desventaja, dado que Fodor no da ningún argumento en favor del revisionismo semántico.[133]

Más aun, la apelación no ya a descripciones asociadas sino a capacidades discriminatorias puede conducir a la creencia de que el factor epistémico es más fuerte en esta teoría que en las teorías descripcionales de la referencia. En el marco de éstas últimas, como vimos, el objeto referido es conocido por el hablante en sus propiedades identificadoras, pero no por ello deja de ser externo al hablante.

132. Como digo más arriba, lo que resulta inadecuado es pretender identificar a las relaciones causales semánticamente pertinentes sobre la base del *ejercicio* de las capacidades discriminatorias individuales.

133. Tal vez se piense que las objeciones de Kripke no se aplican a una teoría como la de Fodor porque las primeras se refieren al proceso de comprensión del lenguaje natural mientras que el dominio de la teoría de Fodor es el pensamiento o lenguaje mental. A esto quiero replicar que las objeciones de Kripke no tienen porqué limitarse al lenguaje natural; como señalé anteriormente, la teoría propuesta por Kripke puede ser interpretada como una teoría que se aplica no sólo al lenguaje natural sino también al lenguaje mental.

En el marco de la teoría de Fodor, el hecho de que el objeto referido sea un objeto *discriminable* (en condiciones ideales), además de exacerbar, como se mencionó anteriormente, el carácter identificador del conocimiento involucrado, trae fuertes reminiscencias verificacionistas y, con ellas, la tendencia general a reducir el objeto en cuestión a condiciones del sujeto -que a su vez involucra la tendencia al antirrealismo metafísico tradicionalmente asociado con el verificacionismo.[134]

Esto último no debe ser interpretado como una acusación drástica; como es sabido, hay muchos tipos de verificacionismo y no hay que confundir los mejores con los peores. Por ello, prefiero hablar de *teorías de sesgo verificacionista*, las cuales creo que comprenden un espectro muy amplio.[135] La teoría de Fodor, con su apelación a capacidades discriminatorias idealizadas, podría decirse que se encuentra en el extremo del espectro que no se encuentra comprometido ontológicamente con el antirrealismo. Sin embargo, creo que

134. El argumento que permite pasar de la defensa de una teoría verificacionista del significado a la suscripción del antirrealismo metafísico es, sin embargo, complejo y su explicitación excede los límites de este trabajo. Véase, por ejemplo, Dummett 1975.

135. En un extremo, es posible situar el verificacionismo crudo de los positivistas lógicos, según el cual el significado de una oración es identificado con el método para su verificación. A continuación, tal vez sea posible ubicar a las teorías que identifican el significado oracional con las condiciones de asertabilidad o aserción justificada. Según Dummett, el representante más acabado de este último tipo de teorías, el significado de una oración como "ayer los obreros hicieron una huelga" no está constituido por el hecho de que los obreros hicieron una huelga el día anterior sino por la posesión, por parte de hablantes e intérpretes, de alguna evidencia en favor de ese hecho -tal como una observación o un recuerdo- que los justifica a afirmar o asertar la oración en cuestión. Desde este punto de vista, los significados oracionales están constituidos por conjuntos de evidencias subjetivas; de ahí que el concepto de condiciones de verdad, que hace referencia a condiciones del mundo, sea reemplazado por el de condiciones de asertabilidad, que alude, en cambio, a condiciones del sujeto. Tanto en el verificacionismo como en la teoría de Dummett, esta internalización del significado conduce a su vez a una internalización del mundo, esto es, al antes mencionado antirrealismo metafísico.

coincide con las otras teorías en un rasgo general, que podría considerarse como distintivo de todas ellas: la idea de que el significado depende, en algún sentido, de nuestras capacidades de reconocimiento, de los criterios o métodos para su reconocimiento con los que podamos contar; de este modo, la cuestión ontológica concerniente a la naturaleza o constitución del significado es identificada con la cuestión gnoseológica concerniente a nuestro conocimiento del mismo.

Es posible considerar que esto último se pone de manifiesto en la aplicación del método de las diferencias a través de los mundos posibles. ¿Qué puede querer decir exactamente que a fin de fijar el contenido de una representación (tipo) es necesario considerar, desde una perspectiva ideal, ciertas posibilidades contrafácticas? No puede significar que el contenido de un caso actual de ese tipo depende ontológicamente del contenido de otros casos meramente posibles. ¿Cómo podría lo actual depender en cuanto a su existencia de posibilidades no actualizadas? Pienso que la tesis en cuestión puede interpretarse como la afirmación de que, a fin de conocer el contenido de una representación actual dada, es necesario tomar en cuenta el contenido de otras representaciones no actualizadas. De este modo, no es el contenido mismo sino nuestro conocimiento del contenido lo que puede pensarse que es determinado mediante la consideración de ciertas posibilidades contrafácticas. En otras palabras, la consideración de los mundos posibles cercanos al nuestro no provee un mecanismo para fijar contenidos sino de una suerte de método de testeo de contenidos que han sido fijados de alguna otra manera. Retomando el ejemplo anterior, si (i) podemos concebir el hecho de que, en condiciones similares a las actuales, las vacas causen casos de VACA aun cuando los caballos en la oscuridad no lo hagan, y (ii) no podemos concebir una situación similar en la que los caballos en la oscuridad causen casos de VACA si las vacas no lo hacen, entonces estaremos en condiciones de formular una hipótesis acerca no del contenido de VACA sino de nuestro estado epistémico respecto de ese contenido, a saber, que sabemos que VACA significa <vaca> y no <caballo en la oscuridad>. Nuevamente, en general, el método de las diferencias no parece servir para determinar el contenido sino nuestra posición epistémica respecto del contenido -en otras palabras, nos permite determinar si estamos

o no en condiciones de identificar el contenido de un concepto con su causa. Ahora bien, el punto es que encontrar una solución al segundo problema no significa haber resuelto el primero, a menos, por supuesto, que se considere, a la manera verificacionista, que uno y otro no son sino el mismo problema. En síntesis, es posible considerar que la teoría de Fodor ofrece una explicación no del mecanismo por el cual nuestros conceptos refieren a objetos y adquieren de ese modo sus significados sino del mecanismo, el método de las diferencias a través de los mundos posibles, por el cual nos es posible a nosotros acceder a los significados de nuestros conceptos. Para seguir con el ejemplo inicial, en la medida en que apliquemos el método de las diferencias del modo antes señalado, estaremos justificados a afirmar que VACA significa <vaca> y no <caballo en la oscuridad>. Pero, a menos que uno sea verificacionista en el sentido general de identificar las cuestiones semánticas con cuestiones gnoseológicas, evidenciales o criteriales, no puede pretenderse que el mecanismo en cuestión sea explicativo de la naturaleza del significado. De ahí que pueda plantearse la siguiente disyunción: o bien se considera que se trata de una teoría verificacionista en el sentido general del término o bien se considera que el objetivo principal de Fodor no es explicar la naturaleza del significado. Como esto último es altamente improbable tratándose de un teórico del significado, daré por sentado que Fodor tiene realmente como objetivo principal explicar la naturaleza del significado, para concluir entonces, en primer lugar, que la teoría propuesta por Fodor, en tanto identifica el mecanismo de fijación del significado con un mecanismo de reconocimiento del mismo, es una teoría de sesgo verificacionista.

Cabe agregar, además, que la ausencia de toda consideración acerca del mecanismo de transmisión de la referencia, como un mecanismo distinto del de fijación, también puede ser considerado un índice del sesgo verificacionista (en el sentido general del término, explicitado anteriormente) de la teoría. En general, las teorías que, a diferencia de las de sesgo verificacionista, intentan acotar el rol de las capacidades epistémicas en la determinación del significado, comienzan por distinguir el mecanismo de fijación de la referencia del mecanismo de transmisión de la misma, para luego reducir o eliminar

por completo la existencia de una dimensión epistémica en el último de ellos -como es el caso de las teorías causales de la referencia de Kripke y Putnam y de la teoría descriptivo-causal de Devitt y Sterelny, examinadas en los capítulos anteriores.

Finalmente, cabe destacar que el sesgo verificacionista constituye un rasgo absolutamente paradójico, puesto que la intención de Fodor, al ofrecer la teoría covariacional, es, como se mencionó anteriormente, ofrecer una teoría puramente causal de la referencia -es decir, un tipo de teoría que no reconoce en el significado dimensión epistémica alguna.

1.3.4. La estrategia de exportación

Si bien hay un sentido claro en el cual la teoría covariacional no constituye, en mi opinión, una teoría puramente causal de la referencia, la naturaleza peculiar del factor epistémico involucrado (no descripcional) así como el complejo rol por él desempeñado (en la delimitación de las condiciones ideales en las que opera el mecanismo causal) no permiten considerar, paradójicamente, que se trata en modo alguno de lo que hemos llamado "sentido" o, en general, "modo de presentación". En la medida en que no se compromete con ninguna otra propiedad semántica que no sea la de referir a un objeto, la teoría covariacional constituye sin duda una teoría de la referencia directa. En tanto tal, no provee solución alguna al antes mencionado problema del sentido; en otras palabras, se le plantean los mismos problemas semánticos tradicionales que a todas las teorías de su tipo -a saber, el problema de la identidad, la paradoja de la denotación, el problema de las oraciones acerca de entidades ficticias, el problema de las oraciones existenciales negativas-, a los que se suma el mencionado problema de la falta de poder explicativo. De acuerdo con esto, una teoría que redunda en la asignación de meros referentes y condiciones veritativas, como es el caso de todas las teorías de la referencia directa, no puede explicar la conducta; como vimos, para hacerlo, es necesario admitir o prescribir la asignación de significados más finamente discriminados. El problema es, precisamente, especificar de qué modo preciso se logra el grado de discriminación en cuestión. Como vimos en el capítulo anterior, hay distintas maneras de hacerlo. La propuesta de Fodor

consiste en complementar la teoría covariacional con una tesis acerca del denominado "significado o contenido estricto".[136] Ahora bien, según Fodor, esta tesis no pertenece estrictamente al terreno de la semántica sino al de la filosofía de la mente: no tiene que ver con la naturaleza del significado sino con el criterio de individualización de los estados mentales. En términos de Fodor:

> *Les propongo ahora contarles una historia acerca de la individualización de las creencias y acerca de cómo las fórmulas subordinadas funcionan en las atribuciones de creencia. La significativa peculiaridad de esta historia es que permite que distintos estados de creencia tengan los mismos contenidos (los mismos objetos proposicionales). El punto de contarles esta historia es que dado que tales casos están permitidos, la proposición de que Yocasta es elegible puede llegar a resultar idéntica a la proposición de que la madre de Edipo es elegible aun cuando creer la primera proposición sea un estado diferente del de creer la otra. Pero, si estas proposiciones pueden ser la misma, entonces no tenemos por el momento ninguna razón para dudar de que "Yocasta" y "la madre de Edipo" sean sinónimos. Lo que implica afirmar que, por lo menos en lo que concierne a los hechos acerca de Edipo, no tenemos ninguna razón para dudar de que la denotación agote el significado.*[137]

Sin entrar en los detalles de su propuesta, la idea central de Fodor es que los estados mentales se individualizan no sólo en función de los significados de las oraciones involucradas (en su terminología, "las fórmulas subordinadas") sino también en función de sus vehículos (por lo general, fórmulas del lenguaje del pensamiento) y sus roles funcionales (esto es, el rol desempeñado en las inferencias o, en otras palabras, el rol causal desempeñado en los procesos mentales, lo

136. Véase la presentación de la polémica internalismo vs. externalismo contenida en el capítulo 2, la cual, como se mencionó en su momento, tiene a Fodor como uno de sus principales protagonistas.
137. Fodor 1990: p. 166 y siguientes hasta 172.

que antes llamamos "rol conceptual" o "significado estricto").[138] De este modo, si bien las oraciones

YOCASTA ES ELEGIBLE (COMO ESPOSA)

y

LA MADRE DE EDIPO ES ELEGIBLE (COMO ESPOSA)

tienen el mismo significado (compuesto a partir de los significados de sus términos, esto es, el individuo Yocasta para ambos términos sujeto), en la medida en que no tienen el mismo vehículo (la fórmula YOCASTA y la fórmula LA MADRE DE EDIPO, respectivamente) ni el mismo rol funcional, dan lugar a distintos estados de creencia: como dejan claramente sentado las famosas tragedias debidas a Sófocles, creer la primera no es lo mismo que creer la segunda.

En mi opinión, esto ilustra claramente la adopción de lo que antes se denominó, a propósito de la teoría de Kripke, "una estrategia de exportación": se exporta a otra área, en este caso, la filosofía de la mente, a la que se supone que compete la atribución de estados mentales a los hablantes, un problema tradicionalmente considerado semántico, a saber, el problema de la diferencia en el valor de verdad de oraciones, las denominadas "adscripciones de actitudes proposicionales", cuyos términos componentes tienen la misma referencia o son extensionalmente equivalentes -esto es, como se recordará, uno de los problemas que motivó la propuesta fregueana de la noción de sentido. En términos del ejemplo anterior, se considera que las razones por las cuales

138. Cabe aclarar que, en algún momento, Fodor intentó definir el significado estricto en términos de la *función* que asigna a cierto *pensamiento* en cierto *contexto* ciertas *condiciones de verdad*, es decir, como una función que va de pensamientos y contextos a condiciones de verdad. De acuerdo con esto, el significado estricto es radicalmente inexpresable porque sólo es significado en potencia: llega a serlo cuando es anclado o fundamentado en un determinado contexto. Véase Fodor 1987.

Edipo cree que Yocasta es elegible

es una oración verdadera mientras que

Edipo cree que la madre de Edipo es elegible

es una oración falsa no son razones semánticas; en otras palabras, se sostiene que el problema de explicar la diferencia de valor veritativo entre las oraciones anteriores no compete a la semántica sino a la disciplina que se ocupa de los estados mentales, esto es, la filosofía de la mente. Ahora bien, la adopción de este tipo de estrategia parece ocultar una decisión arbitraria respecto de lo que constituye el objeto de la semántica. Como señala Devitt:

> *[...] una estrategia de exportación siempre necesita un principio básico por el cual ciertos aspectos puedan ser considerados ajenos a la semántica. Los filósofos de la referencia directa no disponen, por lo general, de un principio tal.*[139]

En el caso particular que nos ocupa, es claro que Fodor no da ningún argumento que fundamente la tesis según la cual una diferencia en el vehículo utilizado y en el rol funcional de los términos no cuenta como una diferencia en el significado expresado. De manera sorprendente, dice que todo es cuestión de intuiciones, ante lo cual no puedo dejar de señalar que si tengo alguna intuición al respecto es justamente la opuesta a la presupuesta por Fodor. Dejando de lado las intuiciones personales, dada la historia de la filosofía del lenguaje contemporánea, es sin duda Fodor quien tiene la carga de la prueba en este punto. Por esta razón, la estrategia de exportación no deja de parecerme totalmente *ad hoc*.

A modo de conclusión del análisis de la teoría de Fodor, quisiera hacer algunos comentarios generales. Con el propósito de solucionar tanto el problema del respecto como el del error, Fodor parece

139. Devitt 1996: p. 182. Los destacados son míos.

querer evitar la dicotomía entre dos tipos distintos de situaciones: situaciones del tipo A, en las que el contenido de un concepto puede ser identificado con su causa y situaciones del tipo B, en las que el contenido es distinto de la causa. Se resiste entonces a mirar hacia atrás, al hipotético bautismo inicial, claramente considerado como una situación de tipo A. Pero el rechazo de la consideración diacrónica de las cadenas causales paga un precio: hace necesario el examen del espectro sincrónico de los mundos posibles. ¿Es ésta una estrategia razonable? Como es obvio, la estrategia adoptada impide que la propuesta constituya una explicación del significado dada exclusivamente en términos de relaciones causales entre objetos del mundo real y representaciones mentales. Según Fodor, las relaciones causales con objetos posibles en mundos posibles son esencialmente pertinentes para la determinación del significado mental. Pero entonces, ¿puede considerarse a esta teoría compatible con los requisitos de una explicación fisicalista? Pienso que no lo es. En mi opinión, intentar proveer una explicación fisicalista del significado mental es tratar de explicar las relaciones causales entre los objetos del mundo actual, por un lado, y los signos mentales, por el otro. Estas relaciones causales actuales son semánticamente primitivas, y, por lo tanto, prioritarias respecto de cualquier verdad acerca de situaciones contrafácticas.

Ahora bien, el intento de especificar las relaciones causales constitutivas del significado puede conducir al planteo de las siguientes preguntas: ¿se trata de relaciones causales directas, como la que existe entre la temperatura y el agua pura? ¿O se trata de relaciones mediatizadas por convenciones? Es plausible considerar, a diferencia de Fodor, que las relaciones entre conceptos y objetos, en su mayor parte, se derivan de las relaciones entre palabras y objetos, y, por lo tanto, no pueden ser concebidas como relaciones directas. Como es obvio, la mera presencia de ciertos objetos no nos conduce a emitir determinados sonidos -o a escribir determinados grafismos. No hay relación directa entre las vacas y los casos de VACA porque no la hay entre las vacas y los casos de "vaca". Sin embargo, es plausible pensar que existe una relación indirecta: si uno pertenece a una comunidad que acordó llamar a las vacas "vacas", usará "vaca" para referirse a las vacas, y, por tanto, tendrá el concepto VACA con el significado

<vaca>. En lo que concierne a los lenguajes naturales, no parece fácil librarse de los nexos históricos o diacrónicos. Y la razón parece ser que son justamente esos nexos los que importan para la determinación del significado: si queremos ser más específicos en la caracterización de las relaciones causales constitutivas del significado, necesitamos remontarnos a las convenciones iniciales transmitidas a lo largo de la historia de una comunidad lingüística.

De este modo, en mi opinión, la mayor parte de las representaciones mentales proceden de la internalización de las representaciones lingüísticas. Los conceptos que podemos tener dependen, en su mayor parte, de las palabras con que contemos para expresarlos -a modo de ejemplo, ¿cómo podría tenerse el concepto TAUTOLOGIA si no se hubiera aprendido la palabra "tautología" en un curso de lógica? La dimensión social del lenguaje proporciona una fuente más amplia para el pensamiento que las capacidades discriminatorias, sean éstas *standard* o idealizadas. Si esto es así, entonces una explicación completa del significado en términos naturalistas o compatibles con el fisicalismo no podrá evitar la consideración de los nexos histórico-causales entre el lenguaje natural y el mundo.

Sin embargo, esto deja intacto nuestro problema central: no hay que olvidar que la explicación en términos de los nexos histórico-causales no puede generalizarse debido al problema del respecto. Y es entonces plausible pensar que el subconjunto de conceptos básicos para los que no puede darse una explicación en esos términos no derivan su significado de las correspondientes expresiones del lenguaje natural sino que, por el contrario, éstas últimas lo obtienen de los primeros. En otras palabras, podemos conceder que la apelación a los nexos histórico-causales es requerida tanto por la semántica del lenguaje natural como por la psicosemántica en la medida en que ésta es parasitaria de la linguosemántica; pero ¿qué ocurre con la psicosemántica en la medida en que (i) no es parasitaria de la linguosemántica o (ii) la linguosemántica es parasitaria de ella? En síntesis, ¿qué ocure con lo que he denominado anteriormente "semántica primaria"? Es en este punto donde se replantea, en toda su pureza y fuerza, el problema de la intencionalidad. Por consiguiente, el desafío sigue siendo lograr ofrecer una explicación causal no histórica para el subconjunto de conceptos básicos. En otras

palabras, el desafío sigue siendo elaborar una semántica primaria en términos puramente causales. En la próxima sección, me centraré entonces en la propuesta de la llamada "biosemántica" o "enfoque teleológico-causal".

2. El enfoque teleológico-causal: la teoría de Millikan

2.1. Una explicación puramente teleológica

El concepto clave en el enfoque puramente teleológico de la referencia es el de función propia. Para Millikan, aquello a lo que una representación refiere y que constituye su contenido o significado está determinado por la función propia de la representación en cuestión. A grandes rasgos, la función propia de algo equivale a su propósito o *télos*, a aquello que hace que el objeto en cuestión pertenezca a cierta categoría biológica -de ahí que la teoría teleológica de Millikan sea también denominada "teoría biosemántica". A continuación, propongo entonces hacer un pequeño *excursus* que nos permita arrojar luz sobre esta noción clave y clarificar en la misma medida la tesis semántica enunciada.

Ante todo, cabe aclarar que Millikan suscribe la definición histórica del concepto de función propia, según la cual la función propia de un ítem determinado es una función que ha sido desempeñada por sus ancestros y que ha ayudado a explicar la proliferación de genes responsables de ese ítem. Esta definición exige entonces: (i) que haya reproducción o copia del rasgo funcional a lo largo de la historia; y (ii) que el rasgo en cuestión haya sido objeto de selección natural. Ambos rasgos son claramente enfatizados por Millikan:

> *En primer lugar, exijo que la manera en la cual el ítem funcional fue producido desde la época de sus ancestros involucre reproducción o copia de sus rasgos funcionales. Mínimamente, la manera en la cual el ancestro es responsable de la existencia presente del ítem funcional debe ser tal que, si aquellas propiedades del ancestro que explican su habilidad para realizar la función hubieran sido*

distintas en ciertos respectos, la progenie habría diferido de manera acorde.[140]

En otras palabras, cuando se elige entre las múltiples funciones de un ítem, aquélla o aquéllas que constituyen sus funciones propias, es necesario tener en cuenta qué es lo que ha incrementado su valor de supervivencia a lo largo de su historia evolutiva. Para tomar un ejemplo formulado originariamente por Dretske, la función propia del magnetosomo, poseído por ciertas bacterias, no es orientar a la bacteria hacia el norte magnético o hacia el fondo del mar (cosas que de hecho hace) sino alejarla de la superficie del agua, cuya alta dosis de oxígeno resulta tóxica para ese organismo.[141] La definición histórica es pues una definición teórica, que pone en el centro de la teoría semántica a los conceptos claves de la teoría de la evolución de Darwin, tales como los de reproducción genética, valor de supervivencia, selección natural. De este modo, desde la perspectiva histórica, el que un ítem tenga o no determinada función propia depende de si ha tenido o no la historia evolutiva apropiada -lo cual impide considerar que, por ejemplo, un doble nuestro surgido en virtud de un accidente cósmico comparta nuestras funciones propias.

Ahora bien, no sólo los órganos, mecanismos y sistemas completos sino también sus estados y actividades tienen funciones propias. A modo de ejemplo, la danza de una abeja tiene la función propia de indicar el lugar donde está el néctar, el estado de un camaleón sentado sobre una roca tiene la función propia de engañar a su depredador, el despliegue del cuello de un lagarto tiene la función propia de espantar a los intrusos. Del mismo modo, no sólo el mecanismo representacional en general sino cada estado particular de ese mecanismo, a lo que Millikan denomina "ícono intencional", tiene una función propia determinada.

La función propia de un ícono intencional constituye su elemento imperativo. El elemento indicativo, por otro lado, está constituido

140. Millikan 1995: p. 34.
141. Véase Dretske 1988: p. 63.

por la condición Normal del mundo para el desarrollo de la función en cuestión.[142] La condición Normal es aquélla que aparece en una explicación Normal, es decir, la explicación que toma en cuenta el ejercicio y desarrollo histórico-evolutivo de la función propia o la que nos dice de qué manera el ítem en cuestión se las ha ingeniado históricamente para desarrollar su función propia.[143] El punto es que es sólo la condición Normal del mundo, y no cualquier otra condición que permita, eventualmente, el desarrollo de la función propia, la que es pertinente para la constitución del elemento indicativo. En síntesis, cada estado mental o ícono intencional tiene dos elementos: un elemento imperativo constituido por su función propia y un elemento indicativo constituido por la condición Normal para el desarrollo de esa función. (Los íconos no intencionales, como la antes mencionada danza de una abeja o el chapoteo de la cola de un castor en el agua, carecen, en cambio, de elemento indicativo).

Los íconos intencionales paradigmáticos son las creencias. El elemento imperativo o función propia de una creencia no es única sino múltiple: ayudar a producir el cumplimiento de ciertos deseos y participar en inferencias. El elemento indicativo está dado entonces por la condición del mundo que debe darse para que cumpla con sus funciones propias de manera Normal y de acuerdo con una explicación Normal. Ahora bien, este elemento indicativo es precisamente lo que constituye el significado o contenido de la creencia en cuestión. De ahí que Millikan afirme:

> *Lo que sostendré es que las creencias son clasificadas (como creencias de que p versus creencias de que q) no de acuerdo con ninguna función particular que puedan tener sino de acuerdo con ciertas condiciones que deben ocurrir si es que van a cumplir con sus funciones propias (por ejemplo, ayudar a satisfacer ciertos deseos) de acuerdo con una explicación Normal.*[144]

142. La convención de escribir "Normal" con mayúscula está tomada de Millikan e indica que el concepto involucrado es teleológico y no estadístico.
143. Véase, por ejemplo, Millikan 1995: pp. 86-7.
144. Millikan 1995: p. 71.

De modo que este breve rodeo nos permite alcanzar la definición buscada: el contenido de un estado mental o ícono intencional, cuyo caso paradigmático es el de las creencias, está determinado por la condición Normal del mundo que permite el desarrollo de sus funciones propias. Una creencia es verdadera si y sólo si se corresponde con cierta condición del mundo, a saber, aquélla que permite que sus funciones propias se desarrollen de manera Normal. Del mismo modo, un concepto refiere a un objeto en virtud de constituir este último la condición del mundo que permite el desarrollo Normal de su función propia; la relación de referencia es teleológico-causal porque es la relación causal con aquel objeto del mundo que hace posible que el concepto en cuestión desarrolle su *télos* o función propia. Véanse, por ejemplo, las siguientes palabras de Millikan:

> *Las representaciones que el cerebro manipula son símbolos -esto es, cosas que tienen una forma significativa- sólo en la medida en que constituyen, en primer lugar, ítems semánticos, ítems que se corresponden (map onto) con el mundo cuando tienen éxito en la realización Normal de sus funciones propias.*[145]

Millikan adhiere explícitamente al principio de composicionalidad cuando afirma que las representaciones (en general, es decir, en tanto sistema representacional) son articuladas: el nivel representacional varía de acuerdo con las variaciones en la situación representada. Esta adhesión le permite, como es usual, dar cuenta de ciertos rasgos del lenguaje, tales como su carácter productivo y sistemático.[146]

La explicación teleológico-causal de la referencia propuesta por Millikan implica entonces una peculiar defensa de la teoría correspondentista de la verdad. Millikan considera que el sistema conceptual en su conjunto constituye un "mapa interno" del mundo; del mismo modo, los distintos estados mentales constituyen mapas

145. Millikan 1995: pp. 81-2.
146. Véase Millikan 1995: p. 90.

internos de distintas condiciones del mundo -así como la danza de las abejas, para utilizar el ejemplo anterior, constituye un mapa de los caminos que conducen al néctar. La teoría involucra por tanto una concepción *pictórica* de la relación entre el pensamiento y el lenguaje, por un lado, y el mundo, por otro, y la clave para entender esta relación pictórica de representación está dada por la historia evolutiva.[147] De esta manera, Millikan considera haber dado una respuesta específica y novedosa a nuestro problema de ambigüedad.

Por último, cabe hacer dos aclaraciones. En primer lugar, hay ciertas consideraciones que parecen contradecirse mutuamente: por un lado, las teorías teleológico-causales, como la de Millikan, fueron caracterizadas, al comienzo de este capítulo, como teorías causales no históricas; por otro, como se explicita repetidamente en este apartado, la teoría de Millikan está basada en la definición *histórica* de función propia. Vale entonces la pena aclarar los distintos sentidos de la expresión "histórico" involucrados en estas afirmaciones. Por un lado, la teoría de Millikan, como toda teoría teleológico-causal de la referencia, no es una teoría histórica porque no define la referencia en términos de la relación histórico-causal entre una representación y un objeto: por el contrario, lo hace, como vimos, en términos de la relación teleológico-causal entre ambos, es decir, la relación causal entre la representación y la condición Normal del mundo para el desarrollo de su función propia. Si la palabra "histórico" alude a la historia de los usos individuales, en distintas comunidades lingüísticas, de las representaciones de un cierto lenguaje natural, entonces la teoría de Millikan no es una teoría semántica histórica. Por otro lado, la teoría en estudio es una teoría histórica en el sentido en que toma en cuenta la historia evolutiva de las representaciones para determinar sus funciones propias y, consiguientemente, sus referentes teleológico-causales. Si "histórico" alude en cambio a la historia evolutiva de las representaciones, la teoría de Millikan es una teoría semántica histórica. De este modo, cabe destacar que en cierto sentido las teorías teleológicas se asemejan

147. Véase, en particular, el artículo denominado "Thoughts without Laws": pp. 77-81.

más a las históricas que a las indicativas o informacionales: en los dos primeros casos, a diferencia de lo que ocurre con las teorías indicativas, las únicas relaciones causales que cuentan son las relaciones causales efectivas, mientras que las meramente posibles no cumplen rol semántico alguno.

En segundo lugar, cabe aclarar que la teoría de Millikan es una teoría de la referencia directa, es decir, una teoría que sólo se compromete con la dimensión referencial o representacional del significado.

2.2. *La solución del problema de los positivos falsos*

En este apartado, explicitaré de qué manera la teoría de Millikan resiste firmemente el embate de la objeción basada en el "problema de los positivos falsos", que hacen tambalear a otras teorías teleológicas.[148] El problema en cuestión puede ser brevemente reconstruido en términos de un ejemplo como el que figura a continuación. El escorpión, que vive sumergido en la arena, necesita, como todo ser vivo, alimentarse; cada movimiento de la arena cercana a él le sugiere la presencia de una presa, a la que intenta manotear; pero sólo uno de cada diez manotazos es recompensado con una presa. Ahora bien, en la medida en que eso es suficiente para mantenerlo alimentado y con vida, la existencia de numerosos intentos fallidos de alimentarse o *positivos falsos* no es grave siempre y cuando haya algunos intentos exitosos. (Grave es, en cambio, que se no dé ningún manotazo cuando de hecho hay una presa en la arena; en otras palabras, a fin de sobrevivir, es importante evitar los negativos falsos.) Sin embargo, en virtud de la existencia de positivos falsos, no es posible afirmar que las instancias de PRESA (en la mente del escorpión), que dependen del funcionamiento óptimo de su mecanismo representacional, covarían o están confiablemente relacionadas con las presas reales; en otras palabras, en los casos de positivos falsos, lo que el concepto PRESA

148. Véase Godffrey-Smith 1992 y Devitt 1991a.

indica (esto es, arena o vacío) no es lo que representa en la mente del escorpión (una presa). Luego, la existencia de correlaciones confiables entre las representaciones y los objetos del mundo, basadas en el cumplimiento de la función propia biológica del mecanismo representacional, no permite fundar hechos semánticos.

Como es claro, este problema da lugar a una grave objeción para las teorías indicativas que incorporan elementos teleológicos, como la teoría de Dretske, y para las teorías teleológicas que se basan en el concepto disposicional de función propia, pero no para teorías como la de Millikan, basada en el concepto histórico de función.[149] Por un lado, una teoría indicativo-teleológica como la de Dretske sostiene que la relación semántica de referencia está constituida por una correlación causal confiable en circunstancias óptimas o Normales, es decir, en aquellas circunstancias en las que el mecanismo de representación cumple con la función propia para la cual fue naturalmente seleccionado, esto es, funciona adecuadamente.[150] Aquí puede apreciarse claramente que el concepto de circunstancias óptimas, concepto de carácter teleológico, es introducido para solucionar el problema del error, esto es, el problema característico de las teorías indicativo-causales: aun cuando de hecho tanto las vacas como los caballos en la oscuridad causen casos de VACA, se supone que en circunstancias óptimas sólo las vacas lo hacen. Ahora bien, para tomar el ejemplo anterior, las circunstancias en las que el escorpión manotea en el vacío son circunstancias en las que su mecanismo de representación cumple con su función propia o funciona adecuadamente; sin embargo, el contenido de la representación instanciada en la mente del escorpión (PRESA) no está dado por lo que la representación en cuestión indica o por aquello con lo que está causalmente correlacionada de manera confiable (nada, o un vacío en la arena). Por lo tanto, parecería que, en virtud de la existencia de

149. Cabe destacar que también parece ser un problema insoluble para la teoría puramente teleológica propuesta por Papineau en Papineau 1987, en la medida en que ésta requiere la verdad de todas las creencias naturalmente seleccionadas; véase al respecto Devitt 1991ba.
150. Véase Dretske 1988.

positivos falsos, la noción de correlación causal confiable en circunstancias óptimas no es suficiente para definir la referencia.

Por otro lado, las teorías teleológicas disposicionales son, a diferencia de la anterior, teorías puramente teleológicas, que atribuyen funciones propias no ya al mecanismo representacional en su conjunto sino a cada uno de sus estados particulares. De acuerdo con esto, la referencia de una representación dada está determinada no ya por la existencia de una correlación confiable en circunstancias de óptimo funcionamiento del sistema representacional sino por la existencia de una relación causal basada en el ejercicio de una función propia de la representación en cuestión. Pero el concepto de función propia es concebido en términos del conjunto de propiedades y disposiciones *presentes*.[151] Ahora bien, los casos de positivos falsos son casos en los que las representaciones particulares involucradas no cumplen con sus funciones propias; en términos del ejemplo anterior, el concepto PRESA en la mente del escorpión no cumple con la función de contribuir a la alimentación del escorpión. Pero si la función propia de un ítem depende enteramente de las propiedades, disposiciones y capacidades que el ítem de hecho tiene, entonces se requiere una explicación de porqué en tales casos casos, pese a tener las requeridas propiedades, disposiciones y capacidades, el ítem involucrado no ejerce su función propia. Además, en la medida en que las funciones propias no se cumplen, las relaciones causales correspondientes con los objetos del mundo que permiten el desarrollo de esas funciones no tienen lugar; por lo tanto, se trata de casos en los que los contenidos representados no pueden estar constituidos por las relaciones causales en cuestión. Se requiere entonces una explicación de cómo las representaciones involucradas adquieren el contenido que tienen.

¿Qué es lo que hace a la teoría de Millikan fuertemente resistente a esta objeción? La clave reside en el concepto de función propia presupuesto, esto es, el concepto *histórico-evolutivo*. Si, a diferencia de lo sostenido por la concepción disposicional, la función propia de un ítem

151. Véanse Wright 1976, Bigelow y Pargetter 1987, Cummins 1975. Para una defensa distinta de la misma concepción histórica defendida por Millikan, véanse Neander 1991a y 1991b.

depende no ya de sus propiedades actuales sino de su historia evolutiva, entonces es perfectamente posible explicar los casos anteriores: los rasgos funcionales que reproducen rasgos ancestrales que han sido objeto de selección natural no tienen porqué estar siempre presentes o, en otras palabras, las funciones propias, históricamente concebidas, pueden no cumplirse en ciertos casos particulares -inclusive, en mucho de ellos, como muestra el ejemplo de los espermatozoides, los cuales sólo en muy pocas oportunidades cumplen con su función propia, a saber, fertilizar un óvulo. Esto es, por lo demás, un resultado esperable, en la medida en que el desarrollo de las funciones depende en parte de sucesos que tienen lugar en la periferia del organismo y más allá de ella. Los contenidos de las representaciones involucradas pueden, no obstante, ser explicados en términos de relaciones teleológico-causales pues, como se mencionó anteriormente, es el correlato ontológico de la manera histórica de ejercer la función propia, esto es, la condición Normal del mundo que permite ese ejercicio, lo que constituye la referencia y, por tanto, el significado o contenido de una representación. Desde el punto de vista semántico, lo importante no es el comportamiento presente del ítem ni el ejercicio presente, comoquiera que sea, de la función propia sino sólo la manera histórica de ejercerla. Para nuestro escorpión, PRESA significa <presa> en virtud de una relación causal histórico-evolutiva, aunque un caso particular (o muchos de ellos) de la representación en cuestión esté de hecho causalmente relacionado con vacío o con arena.

En síntesis, desde el punto de vista de la teoría de Millikan, los casos de positivos falsos son simplemente casos en los que el ítem involucrado tiene una función propia en virtud de la cual queda determinada la relación semántica de referencia entre una representación y un objeto, aun cuando la función en cuestión no sea de hecho ejercida o desarrollada -debido a circunstancias adversas del mundo externo. Volviendo al ejemplo, aun cuando de hecho el escorpión no ejerza su función alimentaria y, por tanto, no haya relación causal presente entre la representación PRESA instanciada en la mente del escorpión y una presa real, PRESA significa <presa> en virtud de la función alimentaria histórica del escorpión y la consiguiente relación causal histórica entre PRESA y las presas. De ahí que Millikan sostenga, como se mencionó anteriormente, que el

concepto histórico de función propia, a diferencia del concepto disposicional, permite explicar el fracaso de las funciones, la existencia de miembros defectuosos de categorías funcionales y la conducta intencional no exitosa.

El análisis del problema de los positivos falsos nos conduce entonces a destacar un rasgo significativo y especialmente ventajoso de la teoría de Millikan: su manera de concebir la relación entre el éxito y la verdad. El reconocimiento de la posibilidad de que las funciones propias no se cumplan da lugar al reconocimiento de la formación de gran número de creencias falsas. De este modo, es posible pensar que los mecanismos humanos de fijación de creencias fijan muchas veces creencias falsas. Asimismo, los casos en los que la función propia se cumple de manera aNormal, es decir, de manera distinta de cómo se cumplió históricamente, son también casos en los que puede haber éxito (en el cumplimiento de la función propia) sin verdad. En otras palabras, según Millikan, el éxito conductual individual no requiere la verdad de las creencias involucradas e incluso la supervivencia de la especie es compatible con la presencia ocasional de creencias falsas.[152] En palabras de Millikan:

Muchos mecanismos biológicos realizan sus funciones propias no regularmente sino tan sólo lo suficientemente seguido. El color protector

152. Vale la pena hacer la siguiente aclaración. Decir que la supervivencia es compatible con la presencia ocasional de creencias falsas no implica negar que la supervivencia requiera en algún grado la verdad de las creencias. La supervivencia de la especie depende de que las funciones propias se cumplan ocasionalmente de manera *Normal*, y, cuando esto ocurre, las creencias involucradas resultan verdaderas. Ahora bien, la supervivencia es sólo un tipo de éxito, a saber, el éxito de la especie en mantenerse en existencia. Hay otros tipos de éxito, como el éxito conductual individual, que en ningún caso requiere la verdad de las creencias involucradas. De ahí que se afirme que la teoría de Millikan no suscribe la tesis de que el éxito requiere siempre la verdad de las creencias, la cual puede entenderse en términos de que no todo tipo de éxito requiere en algún grado la verdad de las creencias -como lo ejemplifica el caso del éxito conductual individual.

que adoptan las crías de la mayor parte de las especies animales es una adaptación transmitida porque ocasionalmente impide que una cría sea devorada; sin embargo, la mayor parte de las crías de estas especies son igualmente devoradas. Del mismo modo, es posible pensar que los mecanismos que fijan las creencias humanas fijan creencias verdaderas no regularmente sino tan sólo lo suficientemente seguido. [...] De acuerdo con ello, tal vez nuestros mecanismos de consumo de creencias hayan sido cuidadosamente diseñados para tolerar una gran proporción de creencias falsas.[153]

De este modo, la teoría de Millikan no suscribe la discutible tesis de que el éxito requiere *siempre* la verdad de las creencias; en otras palabras, no ha de verse en ella una defensa de la teoría correspondentista de la verdad basada en el argumento de que el éxito sólo puede ser explicado en términos de una noción "sustantiva" de verdad como la correspondentista.[154] Esto último puede verse como una ventaja de la teoría, dado el carácter sumamente cuestionable del mencionado argumento.[155]

2.3. En defensa del concepto de función biológica: una respuesta a Fodor

Este apartado está centrado en la crítica realizada por Fodor, la cual puede sintetizarse en los siguientes términos. La apelación al mecanismo de selección natural no permite decidir entre distintas adscripciones de contenido que contienen términos coextensivos: si todos los Fs son Gs, cualquier cosa que haya sido seleccionada para responder a los Fs también lo ha sido para responder a los Gs. Para poner uno de los ejemplos de Fodor, las ranas se alimentan de moscas, pero puede decirse que

153. Millikan 1995: p. 9. La cita es larga pero interesante para ilustrar el punto que intento destacar.
154. Esta tesis ha sido defendida por varios filósofos; véanse especialmente Boyd 1973, Newton-Smith 1978, McMullin 1984 y Putnam 1981.
155. Para una crítica de esta postura, véanse Devitt 1991b *y* Barrio 1998.

la clase de las moscas es para ellas coextensiva con la clase de las moscas o los *fleebees*, donde algo es un *fleebee* si es una mosca o una bala de rifle de aire comprimido (de hecho, las ranas manotean tanto a las moscas como a las balas en cuestión). Dado que las ranas responden a entidades que son o bien moscas o bien *fleebees*, ¿qué nos permite decir entonces que las ranas tienen un detector natural de moscas en lugar de un detector natural de *fleebees*? En otras palabras, si las moscas son, para las ranas, *fleebees*, las ranas han sido seleccionadas para responder tanto a las moscas como a los *fleebees*; por consiguiente, a la luz del criterio otorgado por el mecanismo de selección natural, no hay modo alguno de decidir si son las moscas o los *fleebees* lo que permite a las ranas cumplir con su función alimentaria; luego, no hay modo de decidir si el contenido representado por las ranas es <mosca> o <*fleebee*>. Desde el punto de vista de la teoría de la evolución y su mecanismo propio, a saber, la selección natural, lo importante es, para seguir con el ejemplo, que las ranas cumplan con la función propia de obtener aquello que les proporciona alimento, independientemente de cómo se describa la función en cuestión -como el cazar moscas, *fleebees*, moscas o *fleebees*, etc. En términos de Fodor,

> *La moraleja es, para repetir, que (...) a Darwin no le importa cómo describes los objetos intencionales de las dentelladas de las ranas. Todo lo que importa para la selección es cuántas moscas se ingenia la rana para ingerir como consecuencia de sus dentelladas, y este número resulta ser exactamente el mismo si uno describe la función del mecanismo alimentario con respecto a un mundo poblado de moscas que son, de facto, fleebees que si uno lo hace con respecto a un mundo poblado de fleebees que son, de facto, moscas. (...) A Darwin le importa cuántas moscas te comes, pero no le importa bajo qué descripción lo hagas.*[156]

En otras palabras, si uno se atiene a las funciones propias seleccionadas mediante el mecanismo de selección natural no es posible

156. Fodor 1990: pp. 72-3. Los destacados son del autor. Véanse también las páginas siguientes del mismo artículo.

asignar un contenido unívoco a las representaciones, por cuanto existen múltiples condiciones del mundo correlacionadas con las funciones en cuestión. En términos del ejemplo anterior, tanto la presencia de moscas como de *fleebees* constituyen condiciones del mundo que permiten el desarrollo de la función alimentaria de las ranas; por lo tanto, no hay modo de saber si el contenido representado por las ranas es <mosca> o <*fleebee*> o <mosca o *fleebee*>. De este modo, el concepto de función propia no resulta adecuado para proveer una definición del contenido. La crítica de Fodor expresa entonces la idea de que las relaciones teleológico-causales presentan el mismo problema de ambigüedad que contamina a las relaciones histórico-causales e indicativo-causales. Por consiguiente, así como en el marco de las teorías históricas e indicativas se apeló a distintos recursos (descripciones, dependencia asimétrica) con el objeto de solucionar este problema, un movimiento semejante debería ser llevado a cabo por los defensores del enfoque teleológico. En opinión de Fodor, esto es, sin embargo, imposible: las teorías teleológicas no son capaces de incorporar los requeridos elementos explicativos, pues no pueden dejar de apoyarse de manera única y central en el concepto, explicativamente insuficiente, de función propia biológica.

Ahora bien, considero que esta crítica no hace justicia a la teoría teleológica de Millikan. En mi opinión, no es cierto que las condiciones del mundo correlacionadas causalmente con las representaciones en virtud de sus funciones propias biológicas estén indeterminadas de la manera en que sostiene Fodor. Nótese que los rasgos funcionales son naturalmente seleccionados en virtud de su poder causal: el mecanismo de selección natural opera sobre aquellos rasgos funcionales que causan un incremento en la aptitud de los individuos correspondientes -lo que a su vez causa la proliferación de los genes responsables de esos rasgos y, consiguientemente, la supervivencia de la especie involucrada.[157] La selección de rasgos

157. Vale la pena aclarar que esto no implica que (i) esos rasgos no presenten mutaciones azarosas, como sostiene la teoría de Darwin ni que (ii) otros rasgos que no causan un incremento en la aptitud no sean también seleccionados.

es, por tanto, una cuestión *causal*. Ahora bien, las condiciones Normales del mundo que permiten el desarrollo de los rasgos en cuestión contribuyen a determinar los respectivos rasgos funcionales; de ahí que la descripción de las condiciones forme parte de la descripción de los rasgos -en términos de uno de los ejemplos anteriores, la danza de las abejas cumple con la función de indicar el néctar, donde el néctar es la condición Normal del mundo correlacionada con la danza. Si esto es así, es decir, si las condiciones del mundo correlacionadas con las representaciones son individualizadas en función de su contribución al rol causal de los rasgos funcionales en la historia evolutiva, entonces no hay indeterminación. Por lo tanto, no es posible afirmar que el contenido de las mismas, definido en términos de dichas condiciones, sea ambiguo o no esté unívocamente determinado.

Tal vez esto resulte más claro si se lo aplica al ejemplo de Fodor antes mencionado. Las razones por las que el contenido representado por las ranas es <mosca> y no <*fleebee*> ni el contenido disyuntivo <mosca o *fleebee*> son las siguientes:

(i) las ranas tienen rasgos cuya función es cazar moscas,

(ii) la función de cazar moscas causa un incremento en la aptitud de las ranas,

(iii) en virtud de (ii), la función de cazar moscas es objeto de selección natural,

(iv) en virtud de (iii), las moscas constituyen la condición Normal del mundo para el desarrollo de esa función,

(v) en virtud de (iv), las representaciones de las ranas están teleológico-causalmente relacionadas con moscas -y no con *fleebees* ni con ningún conjunto disyuntivo de entidades.

Esto puede parecer una simplificación excesiva, pero lo que me interesa dejar en claro es la idea central que se expresa a continuación. Las funciones propias y las correspondientes condiciones Normales para su desarrollo pueden ser perfectamente individualizadas en virtud de sus poderes causales, puesto que las primeras son objeto de selección natural, ésta última opera en virtud de los poderes

causales de los rasgos seleccionados y las segundas son definidas en función de las primeras. Por consiguiente, el contenido representacional, definido en términos de funciones propias y condiciones Normales, puede ser unívocamente determinado; en otros términos, las relaciones causales teleológicas no presentan múltiples candidatos como posibles referentes de una representación dada -en este sentido, se diferencian tanto de las relaciones causales históricas como de las relaciones causales indicativas.

A modo de diagnóstico, tal vez sea posible considerar que la razón del supuesto problema de ambigüedad atribuido a la teoría teleológico-causal sea la falta de distinción entre dos tipos distintos de relaciones causales: por un lado, están las relaciones causales entre las representaciones y las condiciones Normales del mundo; por otro, están las relaciones causales entre la presencia de ciertos rasgos funcionales y el incremento de la aptitud. Sólo las primeras son pertinentes para la semántica, puesto que son ésas las que constituyen la relación (semántica) de referencia. Sin embargo, en el contexto de la teoría teleológico-causal, es importante tener en cuenta a las segundas, puesto que son éstas las que otorgan univocidad a las primeras, y de este modo permiten eludir el problema de ambigüedad que aqueja a las otras versiones del enfoque causal. En otras palabras, las relaciones teleológico-causales fundan relaciones semánticas unívocas porque se basan en explicaciones causales -en términos de rasgos funcionales naturalmente seleccionados- perfectamente diferenciables. Retomando nuestro ejemplo, la explicación causal del incremento en la aptitud en términos de la función de cazar moscas no es equivalente a la explicación causal de ese hecho en términos de la función de cazar *fleebees*; por consiguiente, es la relación causal histórico-evolutiva con moscas y no la eventual –o meramente posible- relación causal con *fleebees* la que funda la relación semántica de referencia.

Cabe aclarar entonces que el contenido así definido -esto es, en términos teleológico-causales- puede ser considerado ambiguo o indeterminado sólo si se considera que los poderes causales en general son ambiguos o indeterminados. En otras palabras, la única ambigüedad o indeterminación que puede atribuirse al contenido definido

en los términos anteriores es la heredada de los poderes causales (de los rasgos funcionales seleccionados) a partir de los cuales queda determinado.[158] Pero dudar de la determinación de los poderes causales es dudar de uno de los conceptos básicos, no sólo de la semántica empírica, sino de la ciencia en general. Con esto último lo que pretendo destacar es que, si hay aquí algún problema, no se trata, como en el caso de la ambigüedad que afecta a los otros tipos de teorías causales, de un problema peculiarmente semántico o de la semántica sino de uno más general que afecta a toda la ciencia.

Es interesante destacar que incluso la denominada "ambigüedad quineana" es eliminada: en términos del ejemplo anterior, no es la habilidad para cazar estadios temporales ni partes no separadas de moscas sino la habilidad para cazar *moscas* lo que ha sido objeto de selección natural. De este modo, el concepto de función propia parece ser más apto que el de capacidad discriminatoria, utilizado por las teorías indicativas, para realizar las distinciones semánticas pertinentes.

2.4. La objeción panglossianista y la restricción del alcance de la teoría

Se ha acusado a las teorías teleológicas en general, esto es, no sólo a la teoría de Millikan, de caer en una posición panglossianista o adaptacionista. La expresión "panglossianismo" está basada en el personaje del Doctor Pangloss de Voltaire, quien exhibe una actitud exageradamente optimista ante el mundo por considerar que cada uno de sus rasgos particulares cumple una función determinada que justifica su existencia -incluso, por ejemplo, los dolores de cabeza. Del mismo modo, el adaptacionismo puede definirse como la posición que tiene una visión excesivamente optimista del proceso de evolución y del mecanismo de selección natural que le es propio: desde este punto de vista, la mayor parte de los rasgos observados en los individuos constituyen *adaptaciones*, es decir, son rasgos que han

158. Véase Godffrey-Smith 1989 para una réplica semejante.

sido naturalmente seleccionados para cumplir con la función que de hecho cumplen.

Desde el punto de vista de la crítica que nos ocupa, las teorías teleológicas son panglossianistas fundamentalmente porque presuponen erróneamente que la capacidad cognitiva humana es una adaptación.[159] Más específicamente, su error consiste en ignorar el hecho de que hay muchos rasgos, entre los que se cuenta el sistema cognitivo del hombre, que no son en realidad rasgos adaptativos o adaptaciones sino *exaptaciones*, es decir, rasgos originados azarosamente o como producto contingente (*by product*) de la selección de otros rasgos adaptativos y aplicados posteriormente al cumplimiento de la función que de hecho cumplen; de este modo, no puede decirse de ellos que tengan una función propia genuina. Desde este punto de vista, el sistema cognitivo humano, en tanto exaptación, es un producto contingente del proceso de evolución natural, de la misma manera en que los famosos "*spandrels*" de la catedral veneciana de San Marcos, es decir, los espacios que resultan de la intersección de sus múltiples cúpulas, son un producto contingente del diseño arquitectónico.[160]

A esto es posible responder, por un lado, que no es en absoluto claro que la capacidad cognitiva humana no haya sido seleccionada para cumplir con el rol que de hecho cumple, es decir, que no constituya una adaptación. Si el criterio para distinguir rasgos adaptativos de meras exaptaciones es el incremento de la aptitud (*fitness enhancement*), entonces la capacidad cognitiva parecería ser adaptativa: bien puede ser ese rasgo lo que hace a los seres humanos aptos o, en otras palabras, la supervivencia de la especie humana bien puede deberse a la capacidad cognitiva de los seres humanos.

Por otro lado, es preciso tener en cuenta que la selección natural no sólo opera sobre rasgos funcionales nuevos sino también sobre rasgos viejos a los que asigna nuevas funciones. Por consiguiente, aun cuando se conceda que la capacidad cognitiva humana se haya

159. Véase Gould y Urba 1982; Devitt 1991b: p. 437.
160. Véase Gould y Lewontin 1979: pp. 73-90.

originado como un producto contingente de la selección de otros rasgos -tesis que, repito, no tiene una justificación clara-, esto no implica que posteriomente no haya sido a su vez seleccionada para cumplir con la función que cumple en la actualidad. Este último sería un caso de selección por preservación de rasgos para el cumplimiento de funciones nuevas, es decir, distintas de las que cumplían originariamente. Como dice la propia Millikan,

> *[...] De lejos, la principal tarea de la selección no es la construcción o invención de nada sino el mantenimiento diario del grupo de genes (gene pool). Lo que la selección hace día a día no es alterar los perfiles de las poblaciones sino explusar sistemáticamente la perniciosa chatarra que de lo contrario se acumularía en el grupo de genes debido al cambio azaroso y a una gran cantidad de mutaciones menores.*[161]

La objeción explicitada anteriormente fue hecha fundamentalmente por biólogos. A continuación, voy a referirme a una objeción muy similar a la anterior realizada por filósofos, a saber, aquélla según la cual el carácter panglossianista o adaptacionista de la teoría de Millikan reside no ya en su concepción del sistema cognitivo humano como rasgo adaptativo sino en la tesis de que cada estado particular del sistema en cuestión constituye también una adaptación y tiene, por tanto, una función propia biológica.[162] En otras palabras, de acuerdo con este punto de vista, la teoría de Millikan es panglossianista por considerar que no sólo el mecanismo representacional en general sino cada representación en particular tiene una función propia a cumplir. Desde el punto de vista de esta crítica, la primera afirmación podría llegar a considerarse plausible pero sostener lo segundo es caer en el panglossianismo más burdo. En palabras de Godffrey-Smith:

> *Los rasgos estructurales del aparato visual son productos de una historia evolutiva, una historia de heredada variación en la aptitud.*

161. Millikan 1995: pp. 46-7.
162. Véase, por ejemplo, Godffrey-Smith 1989.

Pero los estados del sistema visual no son la clase de cosa de la que puede decirse que tiene ese tipo de historia.[163]

En mi opinión, esta nueva versión de la objeción panglossianista es seria pero tiene un alcance limitado: obliga no a abandonar la teoría sino a restringirla. En lo que sigue, intentaré explicar esta posición. Antes de ello, quiero aclarar que lo que voy a presentar no es un argumento contundente en favor de la tesis anterior sino tan sólo una posible línea de investigación en favor de la defensa de una teoría teleológica.

Mientras que es plausible pensar que el mecanismo representacional, tanto en la especie humana como en otras especies animales, tiene una función propia que ha sido objeto de selección natural y posee por tanto una historia evolutiva, es sin duda cierto que lo mismo no puede ser dicho acerca de *todos* los estados particulares de ese mecanismo. En particular, no parece plausible sostener esa tesis respecto de estados mentales como las actitudes proposicionales; a modo de ejemplo, mi creencia actual de que la lectura de *Por tu propio bien* me ha revelado ciertos aspectos ocultos de mi infancia no parece tener una historia evolutiva: en primer lugar, porque nadie la ha tenido anteriormente y, en segundo lugar, porque nadie ha podido tenerla anteriormente, debido a que sus componentes indicadores me involucran directamente a mí misma de una manera especial -desde la primera persona. Respecto de tales estados, no es claro que tengan funciones que han sido naturalmente seleccionadas por haber incrementado la aptitud de nuestros ancestros y que explican la existencia actual de las mismas. Por lo demás, la manera en que esas supuestas funciones son descriptas (participar en inferencias o favorecer el cumplimiento de ciertos deseos) es demasiado general: la argumentación en favor de la tesis defendida sería más sólida si se fuera más específica respecto de las funciones propias atribuidas a estados particulares. Además, suponiendo que estados mentales como las actitudes proposicionales tuvieren funciones propias biológicas, no veo claramente de qué manera las condiciones

163. Godffrey-Smith 1989: p. 542.

Normales del mundo que permiten el ejercicio de tales funciones pueden ser producto de los objetos Normales del mundo que permiten el ejercicio de las funciones propias de los componentes suboracionales - recuérdese que una creencia puede ser concebida como un cierto tipo de relación con una oración del lenguaje natural o mental. En primer lugar, porque, en general, no es claro cuáles puedan ser éstas últimas, esto es, las funciones propias de los conceptos y palabras. En segundo lugar, porque las funciones en general no siempre parecen ser composicionales, en el sentido de que la función atribuida a un todo no siempre puede ser considerada un producto de la función atribuida a las partes componentes (la función de un museo de ciencia, divulgar el conocimiento científico, no es producto de la función de cada una de las piezas que la componen, tales como hacer que un cierto líquido se desplace por un tubo o proyectar círculos de colores en una pantalla blanca); y si lo fuere, parecería que no es *sólo* un producto de ellas (en el ejemplo anterior, puede pensarse que se requiere, además, una decisión institucional). Si esto es así, no es claro que las condiciones del mundo identificadas sobre la base de funciones puedan ser composicionales. No es entonces implausible considerar que el carácter composicional de las actitudes proposicionales plantea algunas dificultades a la teoría.

En mi opinión, sin embargo, esto último no implica que la tesis anterior no pueda ser mantenida respecto de *algunos* estados mentales. Más específicamente, considero plausible pensar que algunos estados particulares, los más cercanos a la periferia del individuo y por tanto al mundo que lo rodea, tienen funciones propias que han sido efectivamente objeto de selección natural: es el poseer esos estados representacionales en particular y no meramente la capacidad de representarse al mundo en general lo que incrementó la aptitud individual, permitiendo de esa manera la supervivencia de la especie humana. Si esto es así, no es absurdo pensar que esas representaciones particulares tienen, al igual que el mecanismo representacional, una historia evolutiva. Más aun, considero que éste último sólo la tiene porque hizo posible la instanciación de ciertas representaciones particulares; de lo contrario, ¿cuál sería el interés de poseer un mecanismo representacional?

Cabe aclarar que, en rigor, son los tipos de representaciones particulares, así como los tipos de mecanismos representacionales, los que tienen una historia evolutiva. En relación con este punto, vale la pena destacar que no son los individuos sino las especies las que evolucionan, y éstas últimas han sido concebidas en general como tipos o clases naturales o, más recientemente, como tipos o clases históricas. [164] [165]

Ahora bien, una pregunta se plantea inmediatamente: ¿cuál es el criterio para seleccionar el subconjunto de representaciones(-tipo) en cuestión? Como sugerí anteriormente, considero que se trata de representaciones referidas a estados internos del sujeto o a características salientes del entorno inmediato, esto es, sensaciones, tanto del sentido interno como de los sentidos externos. Respecto de ambos tipos de sensaciones, cabe señalar que el haberlas tenido ha de haber incrementado la aptitud individual en la prehistoria de la humanidad. A modo de ejemplos, el haber sentido dolor ha de haber inducido al hombre a alejarse de la fuente que lo provocaba; el haber distinguido el rojo del verde muy probablemente le haya permitido alimentarse con diversos frutos; el haber sentido miedo ante la presencia de un depredador ha de haberlo inducido a ocultarse de aquél o a huir: si esto es así, se trata de sensaciones que incrementaron la aptitud en tanto dieron lugar a distintas maneras de evitar la muerte. Cabe destacar, además, que esto coincide con la tesis según la cual las distintas fobias constituyen miedos ancestrales, es decir, se trata de miedos del mismo tipo de los que acechaban a nuestros antepasados prehistóricos.[166] Según lo anterior, no sólo las fobias sino todas las sensaciones básicas, tanto internas como externas, serían ancestrales.

164. Según Darwin, la unidad de evolución está constituida por las especies.
165. Para la concepción de las especies como tipos naturales, véase Kitts y Kitts 1979, quienes se oponen a la concepción de aquéllas como individuos (véanse, por ejemplo, Ghiselin 1974 y Hull 1976, 1978). Para una síntesis del debate en torno a las especies que dividió a biólogos y filósofos de la biología, véase Orlando 1995.
166. Véase, por ejemplo, Griffiths (inédito).

Según Sterelny, por ejemplo, los conceptos perceptuales integran módulos o sistemas modulares, cuyo relativo aislamiento y autonomía respecto de otros tipos de estados mentales los hacen muy semejantes a los mecanismos innatos de ciertos animales, como, por ejemplo, el antes mencionado detector de moscas propio de las ranas.[167] Si de éstos puede decirse que tienen una función propia biológica, no habría razones para no decir lo mismo de los primeros. En términos de Sterelny:

> *La historia selectiva nos permite afirmar que las ranas tiene detectores de moscas, y que un punto negro en la retina de una rana es su representación de una mosca porque es producido por un detector de moscas que funciona Normalmente. Del mismo modo, nos permite afirmar que los humanos tienen, por ejemplo, detectores de caras, y que cierta configuración neuronal es la representación que Eric tiene de la cara de Melanie, porque esa representación es producida por un mecanismo modular con una determinada función, el reconocimiento de caras, que opera Normalmente.*[168]

Tal vez pueda pensarse que existen también módulos que agrupan a las sensaciones del sentido interno y módulos para emociones. En general, no es implausible creer que son modulares las representaciones de aspectos del mundo externo o del individuo que puedan haber tenido alguna utilidad para nuestros ancestros -tales como la distinción de colores en el entorno inmediato, el dolor, el miedo, etc.

Por consiguiente, no estoy de acuerdo con el comentario de Godffrey-Smith, mencionado en la cita que figura más arriba: en mi opinión, así como es posible considerar que el aparato visual humano es producto de una historia evolutiva, también es posible considerar que ciertos estados particulares del mismo, como, por ejemplo, la sensación de verde que ahora tengo, también lo son. Es más: lo primero sólo es posible en virtud de lo segundo.

167. En este punto, Sterelny se basa en la tesis fodoriana de la modularidad de la mente (véase Fodor 1983).
168. Sterelny 1990: p. 137.

De este modo, el final de este capítulo nos permite replantear y contestar la pregunta que le dio origen: ¿cuál es el tipo de explicación más adecuado para el subconjunto de representaciones básicas, esto es, aquellas representaciones que, debido al problema del respecto, no pueden ser explicadas en términos histórico-causales? En otras palabras, ¿qué tipo de teoría puede cumplir el rol de semántica primaria y, de ese modo, servir de complemento a la teoría descriptivo-causal antes sugerida? Mi respuesta presente, de carácter programático, es entonces la siguiente: las explicaciones semánticas primarias deben darse en términos teleológico-causales, en la línea general sugerida por Millikan; de este modo, es la apelación al concepto de teleología o función propia lo que hace posible solucionar el mencionado problema de la identificación del subconjunto básico de representaciones. Resta entonces explicitar de qué manera esta explicación teleológico-causal propuesta para los conceptos básicos se articula con la explicación descriptivo-causal antes defendida para la mayor parte del resto del sistema conceptual. La articulación en cuestión será explicitada, a grandes rasgos, en la conclusión de este trabajo.

Conclusión

La prometida conciliación entre la teoría histórico-causal presentada en el segundo capítulo y lo recientemente expresado acerca de la posibilidad de construir una teoría teleológico-causal presupone una cierta concepción de la relación entre el pensamiento y el lenguaje -la cual fue parcialmente explicitada en ocasión del análisis de la teoría de Fodor. Cabe señalar que suscribo plenamente la hipótesis de la existencia de un lenguaje del pensamiento, pero, a diferencia de autores como Fodor, considero que se trata del mismo lenguaje natural que hablamos o de algo muy cercano a éste -muy probablemente el lenguaje del pensamiento contenga también imágenes que son ajenas al lenguaje natural.[169] Desde mi punto de vista, dado que las representaciones mentales o conceptos no son sino palabras internalizadas, no hay razones para creer, a la manera de Fodor, que la habilidad conceptual no esté profundamente relacionada con la habilidad lingüística y dependa en parte de ella. De este modo, como

169. Para la hipótesis de la existencia de un lenguaje del pensamiento o "mentalés", véase fundamentalmente Fodor 1975. Para una defensa de la posibilidad de identificar a aquél con el lenguaje natural, véase Field 1978. Véase también Devitt y Sterelny 1997: cap. 7.

señalé anteriormente, encuentro plausible la tesis de que no todo sino sólo una parte del pensamiento es previa al lenguaje, mientras que el resto se deriva fundamentalmente de aquél. De acuerdo con esto, las representaciones mentales podrían ser clasificadas en dos grupos diferentes. Por un lado, podría pensarse que existe un subconjunto de conceptos básicos, que son independientes del lenguaje, como muestra el hecho de que, en gran parte, compartimos tales conceptos con criaturas no lingüísticas, como los bebés y algunos animales superiores. Por el contrario, el resto de los conceptos tiene, en su mayor parte, un origen lingüístico: los obtenemos al adquirir el uso de ciertos términos del lenguaje natural -tales como "justicia", "tautología" o "gen". Ahora bien, ¿cuál es la justificación para una clasificación como ésta? En este punto, he de confesar que sólo dispongo de una justificación tentativa y programática.

Por un lado, las representaciones mentales básicas parecen poseer contenidos comunes a todos los miembros de la especie humana, e incluso de otras especies animales -algunos conceptos de colores como VERDE y ROJO, de relaciones espaciales como AL LADO y ATRÁS, de sensaciones térmicas como FRIO y CALIENTE, podrían formar parte de ese subconjunto. Los contenidos de tales conceptos básicos dependen de nuestra capacidad para discriminar objetos (en el sentido general del término, que incluye a las propiedades y relaciones) en un entorno común. Ahora bien, dado que toda percepción está teóricamente condicionada, la tarea de especificar exactamente al subconjunto en cuestión se vuelve engorrosa. Por esta razón, considero que la única manera de hacerlo es apelar a lo que es común a todos los seres humanos, independientemente de las diversas teorías acerca del mundo y los distintos sistemas conceptuales que puedan ostentar, a saber, su historia evolutiva. De acuerdo con esto, encuentro plausible pensar que cada representación mental perteneciente al subconjunto básico ha sido naturalmente seleccionada en virtud de su función propia biológica, esto es, un efecto que ha contribuido a la supervivencia de la especie. Para poner un ejemplo simple, si los seres humanos prehistóricos no hubieran sido capaces de discriminar el rojo del verde, probablemente no habrían podido alimentarse debidamente y

la especie humana se habría extinguido; en otras palabras, es plausible creer que los contenidos de nuestras representaciones ROJO y VERDE están determinados en virtud de sus funciones biológicas, concebidas de manera histórica.[170]

Si esto es así, los contenidos de los conceptos pertenecientes al subconjunto básico pueden ser explicados en términos de una teoría teleológica de la referencia del tipo de la propuesta por Millikan. Los significados de las palabras del lenguaje natural que expresan esas representaciones mentales pueden ser, a su vez, explicados en términos de los contenidos de las representaciones mentales correspondientes. Cabe recordar que, como vimos, la apelación a las funciones propias permite eliminar la ambigüedad planteada por el problema del respecto.

Por otro lado, no creo que esta explicación pueda ser extendida a las representaciones mentales que no pertenecen al subconjunto básico. La razón principal en apoyo de ello es que, en mi opinión, como señalé anteriormente, los contenidos de las representaciones mentales no básicas dependen, por lo general, de los significados de las palabras correspondientes del lenguaje natural que son usadas para expresarlas. En nuestras complejas sociedades contemporáneas, el lenguaje natural provee de una base más amplia para el pensamiento que aquélla de la que hemos sido dotados en virtud del proceso evolutivo. Ahora bien, el lenguaje natural es una actividad social; es el producto de distintas actividades desarrolladas en el interior de una sociedad determinada. En consecuencia, considero que la explicación de los significados lingüísticos no puede dejar de considerar el origen y la transmisión de las palabras entre los miembros de esas sociedades concretas en donde se desarrollan los lenguajes naturales. De acuerdo con esto, pienso que un estudio adecuado del

170. Cabe aclarar que el hecho de que tales conceptos puedan llegar a ser compartidos por miembros de otras especies no es un obstáculo para lo explicitado en el texto: si así fuera, se trataría de especies cuya historia evolutiva es, por lo menos, hasta cierto punto, semejante a la de la especie humana.

lenguaje natural debe centrarse en la consideración de los complejos nexos histórico-causales que vinculan a las palabras con los objetos del entorno en que son usadas.

Ahora bien, cabe tener en cuenta que esto no implica, ni siquiera parcialmente, la adopción de una tesis rígida, opuesta a la de Fodor, según la cual el pensamiento está completamente determinado por el lenguaje.[171] Por el contrario, la propuesta anterior intenta dejar abierta la posibilidad de que ciertos conceptos no básicos sean previos y determinantes de las palabras correspondientes. En otros términos, no es mi intención excluir la existencia de aquellos casos en los que, a partir del dominio de un cierto lenguaje y un cierto sistema conceptual, se llega a la adquisición de un concepto, el cual, no sólo no contaba con una palabra correspondiente, sino que da origen a una nueva palabra. De lo contrario, es decir, si se sostiene una tesis rígida de la determinación del pensamiento por el lenguaje, resulta imposible explicar casos del tipo anterior que suelen ser propios de los descubrimientos científicos. Para poner un ejemplo, Newton pudo pensar la ley de gravedad sin contar con la palabra "gravedad", la cual aparece en el vocabulario de los lenguajes europeos sólo a raíz y a partir del pensamiento de Newton. Se trata entonces de un concepto no básico que es previo a la correspondiente palabra del lenguaje natural. En casos como éste, la explicación histórico-causal antes esbozada deberá aplicarse directamente a los conceptos o representaciones mentales involucrados.

Hecha la aclaración anterior, considero que una teoría histórico-(descriptivo)causal de la referencia, en la línea de la propuesta de Devitt y Sterelny pero, como se mencionó anteriormente, con algunas variantes significativas, puede dar cuenta del significado de la mayor parte de las representaciones del lenguaje natural. El significado de la mayor parte de los conceptos será entonces explicado en términos del significado de las correspondientes expresiones naturales. Cabe destacar que, en este nivel, la ambigüedad o indeterminación es eliminada mediante la introducción de descripciones en el

171. Para la defensa de esta tesis, véase paradigmáticamente Whorf 1956.

bautismo inicial, cuya referencia es a su vez fijada en términos de relaciones causales y otras descripciones -y así sucesivamente, hasta llegar a ciertos términos en cuyo bautismo intervienen, junto a la relación causal, sólo términos del subconjunto básico.[172]

En síntesis, en mi opinión, la explicación del significado requiere la combinación de distintos tipos de teorías acerca de la referencia: por un lado, se requiere una teoría teleológico-causal para algunos conceptos básicos del pensamiento; por otro, una teoría histórica, descriptivo-causal, para la mayor parte de las palabras no básicas del lenguaje natural -y para algunos conceptos igualmente no básicos del pensamiento. La primera teoría, esto es, la semántica primaria, es una teoría psicosemántica, mientras que la segunda es casi exclusivamente una teoría linguosemántica.

172. Dado que considero, a diferencia de Devitt y Sterelny, que las descripciones involucradas en el bautismo inicial de un término forman parte de su significado, la afirmación de que en el bautismo inicial de ciertos términos intervienen, junto a la relación causal, términos del subconjunto básico equivale a la afirmación de que el significado de ciertos términos es explicado en función de la relación causal y de ciertos términos básicos. Ahora bien, esta última afirmación no debe entenderse como la tesis de que todos los términos de un lenguaje pueden ser *definidos* en función de un conjunto de términos básicos o atómicos, tesis reductivista propia del positivismo lógico, que ha resultado claramente falsa. La interpretación reductivista puede evitarse si se toma en cuenta que el significado de los términos no básicos está constituido sólo en parte por los términos básicos: la relación causal con objetos del mundo también desempeña un rol constitutivo.

Bibliografía

Alston, William: 1964. *Filosofía del lenguaje*. Madrid, Alianza, 1974.

Aristóteles: *Metafísica,* versión española de Patricio de Azcárate. Madrid, Espasa Calpe, 1993.

Barrio, Eduardo: 1998. *La verdad desestructurada*. Buenos Aires, Eudeba, Enciclopedia Lógica.

Bigelow, J. y Pargetter, R.: 1987. "Functions", en: *The Journal of Philosophy* 84.

Block, Ned. 1986. "Advertisement for a Semantics for Psychology", en P. French, T. Uehling y H. Wettstein (comps.): *Midwest Studies in Philosophy,* vol. X, *Studies in the Philosophy of Mind.* Minneapolis, University of Minnesota Press, 1986.

Boyd, Richard: 1973. "Realism, Underdetermination and Causal Theory of Evidence", en: *Nous* VII.

Burge, Tyler: 1979. "Individualism and the Mental", en P. French, T., Uehling y H. Wettstein (comps.): *Midwest Studies in Philosophy,* vol.IV, *Studies in Metaphysics*. Minneapolis, University of Minnesota Press, 1979b.

— 1986. "Individualism and Psychology", en: *The Philosophical Review* 95.

— 1988. "Individualism and Self-Knowledge", en: *The Journal of Philosophy* 85.

Carnap, Rudolf: 1947. *Meaning and Necessity. A Study in Semantics and Modal Logic*. Chicago, The University of Chicago Press, 1988 (2ª edición).

Comesaña, Juan Manuel: 1998. *Lógica informal, falacias y argumentos filosóficos*. Buenos Aires, Eudeba, Enciclopedia Lógica.

Chomsky, Noam: 1965. *Aspectos de la teoría de la sintaxis*. Madrid, Aguilar, 1970.

— 1968. *Language and Mind*. Harcourt Brace Jovanovich.

Davidson, Donald: 1967. "Truth and Meaning", en: *Inquiries into Truth and Interpretation*. New York, Oxford University Press, 1984.

— 1969. "True to the Facts", en: Davidson 1984.

— 1977. "Reality without Reference", en: Davidson 1984.

— 1990. "The Structure and Content of Truth", en: *The Journal of Philosophy* 87.

Devitt, Michael: 1974. "Singular Terms", en: *The Journal of Philosophy* 71.

— 1981. "Donnellan's Distinction", en P. French, T. Uelhling and H. Wettstein (comps.): *Midwest Studies in Philosophy*, vol.VI, *The Foundations of Analytic Philosophy*. Minneapolis, University of Minnesota Press, 1981.

— 1985. Critical Notice of Evans 1982 *Australasian Journal of Philosophy* 63.

— 1989. "A Narrow Representational Theory of the Mind", en Stuart Silvers (comp.): *Rerepresentation: Readings in the Philosophy of Psychological Representation*. Dordrecht, Kluwer Academic Publishers.

— 1989b. "Against Direct Reference", en: French, Uehling and Wettstein 1989.

— Inédito. "Theories of Reference".

— 1990. "On Removing Puzzles about Belief Ascription", en: *Pacific Philosophical Quaterly* 71.

— 1991a. "Naturalistic Representation" (review of Papineaus's *Reality and Representation*), en: *Brit. J. Phil. Sci.* 42.

— 1991b. *Realism and Truth.* 2ª edición revisada. Oxford, Basil Blackwell.

— 1996. *Coming to Our Senses: a Naturalistic Program for Semantic Localism.* Cambridge, Cambridge University Press.

— 1997. "Afterword", agregado a la nueva edición de *Coming to Our Senses.*

Devitt, Michael y Sterelny, Kim: 1987. *Language and Reality: An Introduction to the Philosophy of Language.* Oxford, Basil Blackwell. 2ª edición revisada: 1997.

Donnellan, Keith: 1966. "Reference and Definite Descriptions", en: *Philosophical Review* 75.

— 1968. "Putting Humpty Dumpty Together Again", en: *Philosophical Review* 77.

— 1972. "Proper Names and Identifying Descriptions", en D. Davidson and G. Harman (comps.): *The Semantics of Natural Languages.* Dordrecht, Reidel, 1972.

Dretske, Fred: 1981. *Knowledge and the Flow of Information.* Cambridge, MA, The MIT Press.

— 1988. *Explaining Behaviour.* Cambridge, MA, The MIT Press.

Dummett, Michael: 1975. "What is a Theory of Meaning?", en Samuel Guttenplan (comp.): *Mind and Language.* Oxford, Clarendon Press, 1975.

Evans, Gareth: 1982. *The Varieties of Reference,* J. McDowell (ed.) Oxford, Clarendon Press.

— 1985. *Collected Papers.* Oxford, Clarendon Press.

Field, Hartry. 1972. "Tarski's Theory of Truth", en: *The Journal of Philosophy* 69.

— 1977. "Logic, Meaning, and Conceptual Role", en: *The Journal of Philosophy* 74, N°7.

— 1978. "Mental Representation", en: *Erkenntnis* 13, N° 1.

Fodor, Jerry: 1975. *The Language of Thought.* New York, Thomas Y. Crowell Company.

— 1978. "Propositional Attitudes", en: *Monist* 61, N° 4.

— 1981. "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology", en: *Representations.* Cambridge, The MIT Press.

— 1982. "Cognitive Science and the Twin-Earth Problem", en: *Notre Dame Journal of Formal Logic,* vol. 23, N° 2.
— 1983. *La modularidad de la mente.* Madrid, Morata, 1986.
— 1985. "Fodor's Guide to Mental Representation: the Intelligent Auntie's Vade-Mecum", en: *Mind,* vol. 94.
— 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* Cambridge, MA, The MIT Press.
— 1990. *A Theory of Content and Other Essays.* Cambridge, MA, The MIT Press.
— 1994. *The Elm and the Expert: Mentalese and Its Semantics.* Cambridge, MA, The MIT Press.
Fodor, Jerry y LePore, Ernest: 1992. *Holism: A Shopper´s Guide.* Oxford, Basil Blackwell.
Frege, Gottlob: 1892a. "Sobre sentido y referencia", en Simpson (comp.): 1973.
— 1892b. "Sobre concepto y objeto", en: *Ensayos de semántica y filosofía de la lógica,* edición, intorducción, traducción y notas de L. M. Valdés Villanueva. Madrid, Tecnos, 1998.
García-Carpintero, Manuel: 1996. *Las palabras, las ideas y las cosas. Una presentación de la filosofía del lenguaje.* Barcelona, Ariel.
Ghiselin, M. T.: 1974. "A Radical Solution to the Species Problem", en: *Systematic Zoology,* vol. 23.
Godffrey-Smith, Peter: 1989. "Misinformation", en: *Canadian Journal of Philosophy* 19.
— 1992. "Indication and Adaptation", en: *Synthese* 92.
— 1996. *Complexity and the Function of Mind in Nature.* Cambridge, Cambridge University Press.
Gould, Stephen J. and Lewontin, R.: 1979. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Program", en: *Proceedings of the Royal Society,* B205, reeditado en Sober (comp.): *Conceptual Issues in Evolutionary Biology: An Anthology.* Cambridge, The MIT Press, 1994.
Griffiths, Paul: Inédito. *Science, Philosophy and the Emotions.*
Haack, Susan: 1978. *Filosofía de las lógicas.* Madrid, Cátedra, 1982.
Harman, Gilbert: 1982. "Conceptual Role Semantics", en: *Notre Dame Journal of Formal Logic,* vol. 23, N° 2.

— 1987. "(Nonsolipsistic) Conceptual Role Semantics", en E. LePore (comp.): *New Directions in Semantics.* London, Academic Press.

Hull, David: 1976. "Are Species Really Individuals?", en: *Systematic Zoology,* vol. 25.

— 1978. "A Matter of Individuality", en: *Philosophy of Science,* vol. 45.

— 1984. "Historical Entities and Historical Narratives", en C. Hookway (comp.): *Minds, Machines and Evolution.* Cambridge, Cambridge University Press.

Kaplan, David: 1978. "Dthat", en French, Uehling and Wettstein 1979.

— 1989. "Demonstratives: An Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and Other Indexicals", en J. Almog, J. Perry and H. Wettstein (comps.): *Themes from Kaplan.* New York, Oxford University Press.

Kenny, Anthony. 1973. *Wittgenstein.* Madrid, Alianza, 1980.

Kim, Jaegwon: 1982. "Psychophysical Supervenience", en: *Philosophical Studies,* vol. 41.

Kitts, D. B. y Kitts, D. J.: 1979. "Biological Species as Natural Kinds", en: *Philosophy of Science,* vol. 46.

Kripke, Saúl: 1972. "Naming and Necessity", en D. Davidson and G. Harman (comps.): *Semantics of Natural Language.* Dordrecht, Reidel.

— 1979. "A Puzzle about Belief", en *Meaning and Use.* Dordrecht, Reidel.

— 1980. *El nombrar y la necesidad,* traducido al español por M. Valdés. México, UNAM, 1985.

— 1982. *Wittgenstein on Rules and Private Language.* Cambridge, Harvard University Press.

Lewis, David: 1986. *On the Plurality of Worlds.* Oxford, Basil Blackwell.

Loar, Brian: 1982. "Conceptual Role and Truth-Conditions", en: *Notre Dame Journal of Formal Logic,* vol. 23, N° 2.

Loewer, Barry. 1982. "The Róle of 'Conceptual Role Semantics' ", en: *Notre Dame Journal of Formal Logic,* vol. 23, N° 2.

Loewer, Barry and LePore, Ernest: 1989. "Dual Aspect Semantics", en Silvers, S. (comp.), 1989.
Loewer, Barry & Rey, Georges (eds.): 1991. *Meaning in Mind: Fodor and his Critics*. Oxford, Basil Blackwell.
McGinn, Colin: 1982. "The Structure of Content", en Woodfield (comp.): *Thought and Object*. Oxford, Clarendon Press.
McMullin. 1984. "A Case for Scientific Realism", en J. Leplin (comp.): *Scientific Realism*. Berkeley, University of California Press.
Mill, John S.: 1867. *A System of Logic*. London, Longmans, Libro I, cap. II.
Millikan, Ruth: 1984. *Language, Thought and Other Biological Categories: New Foundations for Realism*. Cambridge, MA, The MIT Press.
— 1990. "Seismograph Readings for Explaining Behaviour", en: *Philosophy and Phenomenological Research* 50.
— 1995. *White Queen Psychology and Other Essays for Alice*. Cambridge, MA, The MIT Press.
Neander, Karen: 1991a. "Functions as Selected Effects: The Conceptual Analyst's Defense", en: *Philosophy of Science* 58.
— 1991b. "The Teleological Notion of 'Function' ", en: *Australasian Journal of Philosophy* 69.
Newton-Smith: 1978. "The Underdetermination of Theories by Data", en: *Proceedings of the Aristotelian Society*.
Orlando, Eleonora: 1995. "El debate sobre las especies", en *Actas de las V Jornadas de Epistemología e Historia de la Ciencia*. Córdoba, Facultad de Filosofía y Humanidades.
— 1998. "Verdad, significado y realismo". Tesis doctoral inédita.
Papineau, David: 1984. "Representation and Explanation", en: *Philosophy of Science* 51.
— 1987. *Reality and Representation*. Oxford, Basil Blackwell.
— 1991. "Teleology and Mental States", en: *Proceedings of the Aristotelian Society* 65.
Penelas, Federico: Inédito. "Davidson y el holismo: una respuesta a Fodor y LePore".

Platón: *Sofista* en *Diálogos* V, versión española de Ma. Isabel Santa Cruz, Alvaro Vallejo Campos y Néstor Cordero. Madrid, Gredos, 1988.

Putnam, Hilary: 1975. *Mind, Language and Reality: Philosophical Papers*, vol.2. Cambridge, Cambridge University Press.

— 1978. *Meaning and the Moral Sciences*. London, Routledge and Kegan Paul.

— 1981. *Reason, Truth and History*. Cambridge, Cambridge University Press.

— 1988. *Representación y realidad,* versión española de Gabriela Ventureira. Barcelona, Gedisa, 1990.

— 1994. *Words and Life*. Cambridge, Harvard University Press.

Quine, Willard Van Orman: 1951. "Two Dogmas of Empiricism", en: *The Philosophical Review* 60. Reeditado en *From a Logical Point of View*.

— 1953. *From a Logical Point of View*. New York, Harper and Row.

— 1956. "Quantifiers and Propositional Attitudes", en: *The Journal of Philosophy* 53.

— 1960. *Word and Object*. Cambridge, The MIT Press. Cap.2.

Russell, Bertrand: 1905. "On Denoting" *Mind* 14.

— 1911. "Knowledge by Acquaintance and Knowledge by Description", en: *Mysticism and Logic*. London, George Allen and Unwin, 1956.

— 1912. *The Problems of Philosophy*. London, Oxford University Press, 1959. Cap. V.

— 1918. "The Philosophy of Logical Atomism", en R. C. Marsh (ed.): *Logic and Knowledge*. London, George Allen and Unwin, 1956.

Sabatés, Marcelo: 1993. "Internalismo, individualismo y el argumento modal a favor del contenido estrecho", en: *Análisis Filosófico,* vol. 13, N° 2.

Salmon, Nathan: 1982. *Reference and Essence*. Oxford, Basil Blackwell. Cap. 1 y 2.

Searle, John: 1958. "Proper Names", en: *Mind* 67.

Simpson, Thomas Moro: 1964. *Formas lógicas, realidad y significado.* Buenos Aires, Eudeba.

— (comp.): 1973. *Semántica Filosófica: Problemas y discusiones.* Buenos Aires, Siglo XXI.

Stampe, Dennis: 1979. "Towards a Causal Theory of Linguistic Representation", en French, Uehling and Wettstein 1979a.

Sterelny, Kim: 1990. *The Representational Theory of Mind: An Introduction.* Oxford, Basil Blackwell.

Strawson, Peter: 1959. *Individuals: An Essay in Descriptive Metaphysics.* London, Methuen.

Tarski, Alfred: 1935. "The Concept of Truth in Formalized Languages", en: *Logic, Semantics, Metamathematics.* Oxford, Clarendon Press, 1956. Segunda edición revisada: Hackett Publishing Company, 1983.

— 1944. "La concepción semántica de la verdad y los fundamentos de la semántica", versión española de Mario Bunge, Emilio Colombo, Estela Arias y Lilia Fornasari, en Valdés Villanueva, Luis (comp.): *La búsqueda del significado.* Madrid, Tecnos, 1991.

Whorf, Benjamin: 1956. *Language, Thought and Reality. Selected Writings of Benjamin Lee Whorf.* Cambridge, The MIT Press.

Wittgenstein, Ludwig: 1953. *Philosophical Investigations,* versión inglesa de G.E.M. Anscombe, edición de G.E.M. Anscombe y R. Rhees. Oxford, Basil Blackwell.

Wright, Larry: 1973. "Functions", en: *The Philosophical Review* 82.

Esta edición
se terminó de imprimir en
RIPARI S.A.
General J.G. Lemos 248, Buenos Aires
en el mes de agosto de 1999.



Esta colección contiene diversos trabajos que presentan las principales discusiones de la lógica contemporánea, poniendo especial énfasis en los problemas filosóficos que a ella subyacen. La idea es que cada volumen reconstruya el contenido de un concepto central de la mencionada disciplina. En todos los casos, no sólo se intenta describir el estado actual de las discusiones en torno a ese concepto, sino además ofrecer un punto de vista original.

En CONCEPCIONES DE LA REFERENCIA se aborda uno de los aspectos centrales de la relación entre el lenguaje de la lógica y el lenguaje natural: el problema de los significados de las constantes individuales (aquellas expresiones del lenguaje artificial que presuntamente representan individuos) y de las letras predicativas (aquellas expresiones que presuntamente representan clases de individuos). El libro puede verse como una respuesta a fondo al desafío que en nuestros días presentan las corrientes antirrepresentacionalistas. La idea principal que se defiende, de una manera original y muy convincente, es que la relación de referencia debe explicarse en términos mixtos, esto es, haciendo uso tanto de aspectos descriptivos como de relaciones causales. Este volumen contiene, además, una exposición detallada de las distintas concepciones actuales de la referencia. Las teorías del doble factor de Putnam y Block, la explicación descriptivo-causal de Devitt y Sterelny, así como la explicación covariacional de Fodor y el enfoque teleológico-causal de Millikan son analizados con profundidad.




el 0030
ISBN 950-23-0979-0
9 789502 309798